References

[1] Täu Sieskul, B., and Jitapunkul, S. An MIMO parameter estimation model taking into account Ricean fading channel and stochastically uncorrelated signals: I. Inherent Accuracy Limitations. in preparation, (July 2004).

[2] Täu Sieskul, B., and Jitapunkul, S. An MIMO parameter estimation model with Ricean fading channel and stochastically correlatable signals—part I: Cramér-Rao bound. in preparation, (July 2004).

[3] Van Veen, B. D., and Buckley, K. M. Beamforming: a versatile approach to spatial filtering. IEEE Signal Processing Magazine 5 (April 1988) : 4–24.

[4] Krim, H., and Viberg, M. Two decades of array signal processing research: the parametric approach. IEEE Signal Processing Magazine 13 (July 1996) : 97–94.

[5] Godara, L. C. Application of antenna arrays to mobile communications. II. Beamforming and direction-of-arrival considerations. Proceedings of the IEEE 85 (August 1997) : 1195–1245.

[6] Jäntti, T. P. The influence of extended sources on the theoretical performance of the MUSIC and ESPRIT methods: narrow-band sources. IEEE International Conference on Acoustics, Speech, and Signal Processing 1992 (ICASSP-92) 2 (March 1992) : 429–432.

[7] Valaee, S. Array processing for detection and localization of narrowband, wideband and distributed sources. Ph. D. Dissertation, Telecommunications and Signal Processing Laboratory, Department of Electrical Engineering, McGill University, Montreal, Canada, (May 1994).

[8] Trump, T. Estimation of nominal direction of arrival and angular spread using an array of sensors. Technical Report, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, No. IR-S3-SB-9422 (April 1994).

[9] Trump, T., and Ottersten, B. Estimation of nominal direction of arrival and angular spread using an array of sensors. Technical Report, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, No. IR-S3-SB-9607 (April 1996), also appeared in Signal Processing 50 (April 1996) : 57–69.

[10] Zetterberg, P. Mobile cellular communications with base station antenna arrays: spectrum efficiency, algorithms and propagation models. Ph.D. Thesis, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, No. TRITA-S3-SB-9712 (June 1997).

[11] Asztély, D. Spatial models for narrowband signal estimation with antenna arrays. Licentiate Thesis, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, No. TRITA-S3-SB-9728 (November 1997).

[12] Bengtsson, M. Antenna array signal processing for high rank data models. Ph.D. Thesis, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, No. TRITA-S3-SB-9938 (December 1999).

[13] Forssn, U., Karlsson, J., Johannisson, B., Almgren, M., Lotse, F., and Kronestedt, F. Adaptive antenna arrays for GSM900/DCS1800. IEEE Vehicular Technology Conference 1 (June 1994) : 605–609.

[14] Pedersen, K. I., Mogensen, P. E., and Fleury, B. H. A stochastic model of the temporal and azimuthal dispersion seen at the base station in outdoor propagation environments. IEEE Transactions on Vehicular Technology, 49, 2, (March 2000) : 437–447.

[15] Monzingo, R. A., and Miller, T. W. Introduction to Adaptive Array. NY: John Wiley & Sons, 1980.

[16] Haykin, S. (Editor) Array Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[17] Johnson, D. H., and Dudgeon, D. E. Array Signal Processing: Concepts and Techniques. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[18] Hayes, M. H. Statistical Digital Signal Processing and Modeling. NY: John Wiley & Sons, 1996.

[19] Täu Sieskul, B., and Jitapunkul, S. Towards Laplacian angle deviation model for spatially distributed source localization. accepted to participate in International Symposium on Communications and Information Technologies 2004 (ISCIT 2004)

[20] Ottersten, B., Viberg, M., Stoica, P., and Nehorai, A. Exact and large sample maximum likelihood techniques for parameter estimation and detection in

array processing. Technical Report, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, No. TRITA-SB-9302 (February 1993), also appeared in Radar Array Signal Processing Litva, L., Haykin, S., and Shephred, T. J. (Eds.), Berlin, Germany: Springer Verlag, ch. 4 (1993) : 99–151.

[21] Rao, C. R. Linear statistical inference and its applications. second corrected printing, USA: John Wiley & Sons, 1968.

[22] Muirhead, R. J. Aspect of Multivariate Statistical Theory. NY: John Wiley & Sons, 1982.

[23] Anderson, T. W. An Introduction to Multivariate Statistical Analysis. 2nd edition, NY: John Wiley & Sons, 1984.

[24] Wu, Q., Wong, K. M., Meng Y., and Read, W. DOA estimation of point and scattered sources—vec-MUSIC. IEEE Seventh SP Workshop on Statistical Signal and Array Processing (June 1994) : 365–368.

[25] Valee, S., Champange, B., and Kabal, P., Parametric localization of distributed sources. IEEE Transactions on Signal Processing 43 (September 1995) : 2144–2153.

[26] Shahbazpanahi, S., Valee, S., and Bastani, M. H., Distributed source localization using ESPRIT algorithm. IEEE Transactions on Signal Processing 49 (October 2001) : 2169–2178.

[27] Meng, Y. Stoica, P., and Wong, K. M. Estimation of the directions of arrival of spatially dispersed signals in array processing. Proc. Inst. Elect. Eng., Radar, Sonar, Navigat 143 (February 1996) : 1–9.

[28] Lee, S. R., Choi, J., Song, I., and Lee, Y. U. Distributed source modelling and direction-of-arrival estimation techniques. IEEE Transactions on Signal Processing 45 (April 1997) : 960–969.

[29] Bengtsson, M., and Ottersten, B. Low-complexity estimators for distributed sources. IEEE Transactions on Signal Processing 48 (August 2000) : 2185–2194.

[30] Bengtsson, M., and Ottersten, B. A generalization of weighted subspace fitting to full rank models. Technical Report, Royal Institute of Technology, Stockholm, Sweden, No. IR-S3-SB-0103 (January 2001), also in Transactions on Signal Processing 49 (May 2001) : 1002-1012.

[31] Boujemaa, H., and Marcos, S. Joint estimation of direction of arrival and angular spread using the knowledge of the power angular density. The 13th IEEE International symposium on Personal, Indoor and Mobile Radio Communications 4 (September 2002) : 1517–1521.

[32] Sorelius, J. Subspace-based parameter estimation problems in signal processing. Ph.D. Thesis, Department of Information Technology, Uppsala University, Uppsala, Sweden, (February 1999).

[33] Abdellatif, T., Larzabal P., and Clergeot, H. Linear prediction methods for scattered sources. The 50th IEEE Vehicular Technology Conference 1 (September 1998) : 112–116.

[34] Raich, R., Goldberg, J., and Messer, H. Bearing estimation for a distributed source via the conventional beamformer. Proceedings of Statistical Signal Array Processing Workshop (September 1998) : 5–8.

[35] Ottersten, B., Stoica, P., and Roy, R. Covariance matching estimation techniques for array signal processing applications. Technical Report, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, No. IR-S3-SB-9849 (July 1998), also appeared in Digital Signal Processing, A Review Journal 8 (July 1998) 185–210.

[36] Kay, S. M. Fundamentals of Statistical Signal Processing: Estimation Theory. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[37] Besson, O., and Stoica, P. Decoupled estimation of DOA and angular spread for a spatially distributed source. IEEE Transactions on Signal Processing 48 (July 2000) : 1872–1882.

[38] Besson, O., Vincent, F., Stoica, P., and Gershman, A. B. Approximate maximum likelihood estimators for array processing in multiplicative noise environments. IEEE Transactions on Signal Processing 48 (September 2000) : 2506–2518.

[39] Besson, O., Stoica, P., and Gershman, A. B. Simple and accurate direction of arrival estimator in the case of imperfect spatial coherence. IEEE Transactions on Signal Processing 49 (April 2001) : 730–737.

[40] Ghogho, M., Besson, O., and Swami, A. Estimation of directions of arrival of multiple scattered sources. IEEE Transactions on Signal Processing 49 (November 2001) : 2467–2480.

[41] Van Trees, H. L. Detection, Estimation, and Modulation Theory, Part I Detection, Estimation, and Linear Modulation Theory. USA: John Wiley & Sons, 1968.

[42] Stoica, P. and Moses, R. L. Introduction to spectral analysis. Upper Saddle River, NJ: Prectice-Hall, 1997.

[43] Miller, K. S. Complex Stochastic Processes: An Introduction to Theory and Application. Addison Wesley, 1974.

[44] Jansson, M. On subspace methods in system identification and sensor array signal processing. Ph.D. Thesis, Royal Institute of Technology, Stockholm, Sweden, No. TRITA-REG-9701 (October 1997).

[45] Göransson, B. On parametric methods for source Localization. Ph.D. Thesis, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, No. TRITA-S3-SB-9727 (1997).

[46] Van Trees, H. L. Optimum Array Processing. Part VI of Detection, Estimation, and Modulation Theory. USA: John Wiley & Sons, 2002.

[47] Linebarger, D. A., and Johnson, D. H. The effect of spatial averaging on spatial correlation matrices in the presence of coherent signals. IEEE Transactions on Signal Processing 38 (May 1990) : 880–884,.

[48] Li, H., Stoica, P., and Li, J. Computationally efficient maximum likelihood estimation of structured covariance matrices. IEEE Transactions on Signal Processing 47 (May 1999) : 1314–1323.

[49] Turkington, D. A. Matrix Calculus and Zero-One Matrices: Statistical and Econometric Applications. Cambridge University Press, 2002.

[50] Stoica, P., Larsson, E. G., and Gershman, A. B. The stochastic CRB for array processing: a textbook derivation. IEEE Signal Processing Letters 8 (May 2001) : 148–150.

[51] Magnus, J. R., and Neudecker, H. Matrix Differential Calculus with Applications in Statistics and Econometrics. 2nd Edition, NY: John Wiley & Sons, Inc., 1999.

[52] Anderson, T. W. The Statistical Analysis of Time Series. NY: John Wiley & Sons, Inc., 1971.

[53] Ottersten, B., and Ljung, L. Asymptotic results for sensor array processing. International Conference on Acoustics, Speech, and Signal Processing (ICASSP–89) 4 (May 1989) : 2266–2269.

[54] Ljung, L. System Identification: Theory for the User. 2nd edition, Upper sadler River, NJ: Prentic-Hall, Inc., 1999.

[55] Hogg, R. V., and Craig, A. T. Introduction to Mathematical Statistics. 5th edition, Englewodd Cliffs, NJ: Prentice Hall, Inc., 1985.

[56] Serfling, R. J. Approximation Theorems of Mathematical Statistics. NY: John Wiley & Sons, Inc., 1980.

[57] Viberg, M., Stoica, P., and Ottersten, B. Maximum likelihood array processing in spatially correlated noise fields using parameterized signals. Technical Report, Royal Institute of Technology, Sweden, No. IR-S3-SB-9418, (1994), also appeared in IEEE Transactions on Signal Processing 45 (April 1997) : 996–1004.

[58] Monakov, A., and Besson, O. Direction finding for an extended target with possibly non-symmetric spatial spectrum. IEEE Transactions on Signal Processing 52 (January 2004) : 283–287.

[59] Janssen, P. H. M., and Stoica, P. On the expectation of the product of four matrix-valued Gaussian random variables. IEEE Transactions on Automatic Control 33 (September 1988) : 867–870.

[60] Haykin, S. Adaptive Filter Theory. 3rd edition, Englewood Cliffs, NJ: Prentice-Hall, 1996.

Appendix

Proof of Equation (3.42)

Recall the following relations $\lceil \mathbf{AB} \rfloor = \upsilon_{\mathsf{c}}^{\mathsf{H}}\left(\mathbf{A}^{\mathsf{H}}\right)\upsilon_{\mathsf{c}}\left(\mathbf{B}\right)$ and $\upsilon_{\mathsf{c}}\left(\mathbf{ABC}\right) = (\mathbf{C}^{\mathsf{T}} \otimes \mathbf{A})\upsilon_{\mathsf{c}}\left(\mathbf{B}\right)$. Employing (3.20), we obtain the first and the second derivatives as following:

$$
\begin{aligned}
\frac{\partial}{\partial \theta_n}\ell_{\mathrm{AML}}^{[N_T]}(\boldsymbol{\theta}_\omega) &= \lceil \dot{\widehat{\boldsymbol{\Sigma}}}_{xx}(\theta_n)\widehat{\boldsymbol{\Sigma}}_{xx}^{-1} \rfloor - \lceil \widehat{\boldsymbol{\Sigma}}_{xx}^{-1}\dot{\widehat{\boldsymbol{\Sigma}}}_{xx}(\theta_n)\widehat{\boldsymbol{\Sigma}}_{xx}^{-1}\hat{\boldsymbol{\Sigma}}_{xx} \rfloor \\
&= \lceil (\widehat{\boldsymbol{\Sigma}}_{xx}^{-1} - \widehat{\boldsymbol{\Sigma}}_{xx}^{-1}\hat{\boldsymbol{\Sigma}}_{xx}\widehat{\boldsymbol{\Sigma}}_{xx}^{-1})\dot{\widehat{\boldsymbol{\Sigma}}}_{xx}(\theta_n) \rfloor \\
&= \upsilon_{\mathsf{c}}^{\mathsf{H}}(\dot{\widehat{\boldsymbol{\Sigma}}}_{xx}(\theta_n))\upsilon_{\mathsf{c}}(\widehat{\boldsymbol{\Sigma}}_{xx}^{-1}(\widehat{\boldsymbol{\Sigma}}_{xx} - \hat{\boldsymbol{\Sigma}}_{xx})\widehat{\boldsymbol{\Sigma}}_{xx}^{-1}) \\
&= \dot{\widehat{\boldsymbol{\xi}}}_x^{\mathsf{H}}(\theta_n)\widehat{\boldsymbol{\Psi}}_{xx}^{-1}(\widehat{\boldsymbol{\xi}}_x - \hat{\boldsymbol{\xi}}_x)
\end{aligned}
\tag{1}
$$

and

$$
\begin{aligned}
\frac{\partial^2}{\partial \theta_n \partial \theta_{\acute{n}}}\ell_{\mathrm{AML}}^{[N_T]}(\boldsymbol{\theta}_\omega) &= \lceil \ddot{\widehat{\boldsymbol{\Sigma}}}_{xx}(\theta_n, \theta_{\acute{n}})\widehat{\boldsymbol{\Sigma}}_{xx}^{-1} \rfloor \\
&\quad - \lceil \ddot{\widehat{\boldsymbol{\Sigma}}}_{xx}(\theta_n, \theta_{\acute{n}})\widehat{\boldsymbol{\Sigma}}_{xx}^{-1}\widehat{\boldsymbol{\Sigma}}_{xx}^{-1}\hat{\boldsymbol{\Sigma}}_{xx} \rfloor \\
&\quad - \lceil \widehat{\boldsymbol{\Sigma}}_{xx}^{-1}\dot{\widehat{\boldsymbol{\Sigma}}}_{xx}(\theta_n)\widehat{\boldsymbol{\Sigma}}_{xx}^{-1}\dot{\widehat{\boldsymbol{\Sigma}}}_{xx}(\theta_{\acute{n}}) \rfloor \\
&\quad + \lceil \widehat{\boldsymbol{\Sigma}}_{xx}^{-1}\dot{\widehat{\boldsymbol{\Sigma}}}_{xx}(\theta_n)\widehat{\boldsymbol{\Sigma}}_{xx}^{-1}\dot{\widehat{\boldsymbol{\Sigma}}}_{xx}(\theta_{\acute{n}})\widehat{\boldsymbol{\Sigma}}_{xx}^{-1}\hat{\boldsymbol{\Sigma}}_{xx} \rfloor \\
&\quad + \lceil \widehat{\boldsymbol{\Sigma}}_{xx}^{-1}\dot{\widehat{\boldsymbol{\Sigma}}}_{xx}(\theta_{\acute{n}})\widehat{\boldsymbol{\Sigma}}_{xx}^{-1}\dot{\widehat{\boldsymbol{\Sigma}}}_{xx}(\theta_n)\widehat{\boldsymbol{\Sigma}}_{xx}^{-1}\hat{\boldsymbol{\Sigma}}_{xx} \rfloor.
\end{aligned}
\tag{2}
$$

Observe that $\dot{\widehat{\boldsymbol{\xi}}}_x(\theta_n) = \dot{\boldsymbol{\Omega}}(\theta_n)\widehat{\boldsymbol{\eta}} + \boldsymbol{\Omega}\dot{\widehat{\boldsymbol{\eta}}}(\theta_n)$ requires $\dot{\widehat{\boldsymbol{\eta}}}(\theta_n)$ according to

$$
\begin{aligned}
\dot{\widehat{\boldsymbol{\eta}}}(\theta_n) &= \left(\frac{\partial}{\partial \theta_n}(\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\right)\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\hat{\boldsymbol{\xi}}_x + (\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\left(\frac{\partial}{\partial \theta_n}\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\hat{\boldsymbol{\xi}}_x\right) \\
&\overset{(3.20c)}{=\!=} -\left((\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\left(\frac{\partial}{\partial \theta_n}(\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})\right)(\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\right)\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\hat{\boldsymbol{\xi}}_x \\
&\quad + (\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\dot{\boldsymbol{\Omega}}^{\mathsf{H}}(\theta_n)\hat{\boldsymbol{\Psi}}_{xx}^{-1}\hat{\boldsymbol{\xi}}_x \\
&= -\left((\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\left(\dot{\boldsymbol{\Omega}}^{\mathsf{H}}(\theta_n)\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega} + \boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\dot{\boldsymbol{\Omega}}(\theta_n)\right)(\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\right)\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\hat{\boldsymbol{\xi}}_x \\
&\quad + (\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\dot{\boldsymbol{\Omega}}^{\mathsf{H}}(\theta_n)\hat{\boldsymbol{\Psi}}_{xx}^{-1}\hat{\boldsymbol{\xi}}_x \\
&= -(\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\dot{\boldsymbol{\Omega}}^{\mathsf{H}}(\theta_n)\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega}(\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\hat{\boldsymbol{\xi}}_x \\
&\quad - (\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\dot{\boldsymbol{\Omega}}(\theta_n)(\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\hat{\boldsymbol{\xi}}_x \\
&\quad + (\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\dot{\boldsymbol{\Omega}}^{\mathsf{H}}(\theta_n)\hat{\boldsymbol{\Psi}}_{xx}^{-1}\hat{\boldsymbol{\xi}}_x \\
&\overset{(3.24)\&(3.25)}{=\!=} -(\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\dot{\boldsymbol{\Omega}}^{\mathsf{H}}(\theta_n)\hat{\boldsymbol{\Psi}}_{xx}^{-1}\widehat{\boldsymbol{\xi}}_x - (\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\dot{\boldsymbol{\Omega}}(\theta_n)\widehat{\boldsymbol{\eta}}_{\mathrm{AML}}(\boldsymbol{\vartheta}_\omega) \\
&\quad + (\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\dot{\boldsymbol{\Omega}}^{\mathsf{H}}(\theta_n)\hat{\boldsymbol{\Psi}}_{xx}^{-1}\hat{\boldsymbol{\xi}}_x \\
&= (\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\dot{\boldsymbol{\Omega}}^{\mathsf{H}}(\theta_n)\hat{\boldsymbol{\Psi}}_{xx}^{-1}(\hat{\boldsymbol{\xi}}_x - \widehat{\boldsymbol{\xi}}_x) - (\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^{\mathsf{H}}\hat{\boldsymbol{\Psi}}_{xx}^{-1}\dot{\boldsymbol{\Omega}}(\theta_n)\widehat{\boldsymbol{\eta}}_{\mathrm{AML}}(\boldsymbol{\vartheta}_\omega).
\end{aligned}
\tag{3}
$$

In asymptotic region, one can achieve

$$\begin{aligned}\bar{\dot{\hat{\boldsymbol{\eta}}}}_{\text{AML}}(\theta_n) &\triangleq \lim_{N_T\to\infty} \mathcal{E}\big\langle \dot{\hat{\boldsymbol{\eta}}}_{\text{AML}}(\theta_n)\big\rangle \\ &= -\big(\boldsymbol{\Omega}^{\mathsf{H}}\boldsymbol{\Psi}_{xx}^{-1}\boldsymbol{\Omega}\big)^{-1}\boldsymbol{\Omega}^{\mathsf{H}}\boldsymbol{\Psi}_{xx}^{-1}\dot{\boldsymbol{\Omega}}(\theta_n)\boldsymbol{\eta} \\ &= -\big(\boldsymbol{\Omega}^{\mathsf{H}}\boldsymbol{\Psi}_{xx}^{-1}\boldsymbol{\Omega}\big)^{-1}\boldsymbol{\Omega}^{\mathsf{H}}\boldsymbol{\Psi}_{xx}^{-1}\dot{\boldsymbol{\xi}}_x(\theta_n).\end{aligned} \tag{4}$$

Then, the asymptotic derivative $\bar{\dot{\hat{\boldsymbol{\xi}}}}_x(\theta_n) \triangleq \lim_{N_T\to\infty} \mathcal{E}\big\langle \dot{\hat{\boldsymbol{\xi}}}_x(\theta_n)\big\rangle$ is

$$\begin{aligned}\bar{\dot{\hat{\boldsymbol{\xi}}}}_x(\theta_n) &= \dot{\boldsymbol{\Omega}}(\theta_n)\boldsymbol{\eta} - \boldsymbol{\Omega}\big(\boldsymbol{\Omega}^{\mathsf{H}}\boldsymbol{\Psi}_{xx}^{-1}\boldsymbol{\Omega}\big)^{-1}\boldsymbol{\Omega}^{\mathsf{H}}\boldsymbol{\Psi}_{xx}^{-1}\dot{\boldsymbol{\xi}}_x(\theta_n) \\ &= (\boldsymbol{I} - \boldsymbol{\Omega}(\boldsymbol{\Omega}^{\mathsf{H}}\boldsymbol{\Psi}_{xx}^{-1}\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^{\mathsf{H}}\boldsymbol{\Psi}_{xx}^{-1})\dot{\boldsymbol{\xi}}_x(\theta_n) \\ &= \boldsymbol{\Psi}_{xx}^{\frac{1}{2}}\boldsymbol{\Pi}^{\perp}_{\mathcal{S}_{\mathcal{R}}\left(\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\boldsymbol{\Omega}\right)}\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\dot{\boldsymbol{\xi}}_x(\theta_n).\end{aligned} \tag{5}$$

The $(n, \acute{n})$-th element of asymptotic Hessian matrix is then given by

$$\begin{aligned}\big[\bar{\boldsymbol{H}}_{\text{AML}}(\boldsymbol{\theta})\big]_{[n,\acute{n}]} &= \lim_{N_T\to\infty} \mathcal{E}\left\langle \frac{\partial^2}{\partial\theta_n\partial\theta_{\acute{n}}}\ell_{\text{AML}}^{[N_T]}(\boldsymbol{\theta}_\omega)\right\rangle \\ &= \big\lceil \boldsymbol{\Sigma}_{xx}^{-1}\bar{\dot{\hat{\boldsymbol{\Sigma}}}}_{xx}(\theta_n)\boldsymbol{\Sigma}_{xx}^{-1}\bar{\dot{\hat{\boldsymbol{\Sigma}}}}_{xx}(\theta_{\acute{n}})\big\rfloor \\ &= \boldsymbol{v}_{\mathsf{c}}^{\mathsf{H}}(\boldsymbol{\Sigma}_{xx}^{-1}\bar{\dot{\hat{\boldsymbol{\Sigma}}}}_{xx}(\theta_n)\boldsymbol{\Sigma}_{xx}^{-1})\boldsymbol{v}_{\mathsf{c}}(\bar{\dot{\hat{\boldsymbol{\Sigma}}}}_{xx}(\theta_{\acute{n}})) \\ &= \bar{\dot{\hat{\boldsymbol{\xi}}}}_x^{\mathsf{H}}(\theta_n)\boldsymbol{\Psi}_{xx}^{-1}\bar{\dot{\hat{\boldsymbol{\xi}}}}_x(\theta_{\acute{n}}).\end{aligned} \tag{6}$$

Substituting (5) into (6), one can achieve

$$\big[\bar{\boldsymbol{H}}_{\text{AML}}(\boldsymbol{\theta})\big]_{[n,\acute{n}]} = \dot{\boldsymbol{\xi}}_x^{\mathsf{H}}(\theta_n)\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\boldsymbol{\Pi}^{\perp}_{\mathcal{S}_{\mathcal{R}}\left(\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\boldsymbol{\Omega}\right)}\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\dot{\boldsymbol{\xi}}_x(\theta_{\acute{n}}). \tag{7}$$

Straightforward substituting (1) into $\dot{\ell}_{\text{AML}}(\theta_n)\dot{\ell}_{\text{AML}}(\theta_{\acute{n}})$, it results in

$$\dot{\ell}_{\text{AML}}(\theta_n)\dot{\ell}_{\text{AML}}(\theta_{\acute{n}}) = \dot{\hat{\boldsymbol{\xi}}}_x^{\mathsf{H}}(\theta_n)\widehat{\boldsymbol{\Psi}}_{xx}^{-1}(\widehat{\boldsymbol{\xi}}_x - \hat{\boldsymbol{\xi}}_x)(\widehat{\boldsymbol{\xi}}_x - \hat{\boldsymbol{\xi}}_x)^{\mathsf{H}}\widehat{\boldsymbol{\Psi}}_{xx}^{-1}\dot{\hat{\boldsymbol{\xi}}}_x(\theta_{\acute{n}}). \tag{8}$$

Under the same line, the $(n, \acute{n})$-th element of $\big[\overline{\boldsymbol{Q}}_{\text{AML}}(\boldsymbol{\theta})\big]_{[n,\acute{n}]}$ can be written as

$$\begin{aligned}\big[\overline{\boldsymbol{Q}}_{\text{AML}}(\boldsymbol{\theta})\big]_{[n,\acute{n}]} &= \lim_{N_T\to\infty} N_T\mathcal{E}\big\langle \dot{\ell}_{\text{AML}}(\theta_n)\dot{\ell}_{\text{AML}}(\theta_{\acute{n}})\big\rangle \\ &= \dot{\boldsymbol{\xi}}_x^{\mathsf{H}}(\theta_n)\boldsymbol{\Psi}_{xx}^{-1}\left(\lim_{N_T\to\infty} N_T\mathcal{E}\big\langle(\widehat{\boldsymbol{\xi}}_x - \hat{\boldsymbol{\xi}}_x)(\widehat{\boldsymbol{\xi}}_x - \hat{\boldsymbol{\xi}}_x)^{\mathsf{H}}\big\rangle\right)\boldsymbol{\Psi}_{xx}^{-1}\bar{\dot{\hat{\boldsymbol{\xi}}}}_x(\theta_{\acute{n}}).\end{aligned} \tag{9}$$

Without affecting the asymptotic performance, it might be approximated as

$$\big[\overline{\boldsymbol{Q}}_{\text{AML}}(\boldsymbol{\theta})\big]_{[n,\acute{n}]} \simeq \dot{\boldsymbol{\xi}}_x^{\mathsf{H}}(\theta_n)\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\boldsymbol{\Pi}^{\perp}_{\mathcal{S}_{\mathcal{R}}\left(\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\boldsymbol{\Omega}\right)}\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\overline{\boldsymbol{\Sigma}}_{\tilde{\xi}_x\tilde{\xi}_x}\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\boldsymbol{\Pi}^{\perp}_{\mathcal{S}_{\mathcal{R}}\left(\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\boldsymbol{\Omega}\right)}\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\dot{\boldsymbol{\xi}}_x(\theta_{\acute{n}}) \tag{10}$$

where the residual vector $\tilde{\boldsymbol{\xi}}_x \in \mathbb{C}^{N_E^2 \times 1}$ and its limiting covariance $\overline{\boldsymbol{\Sigma}}_{\tilde{\xi}_x\tilde{\xi}_x}$ are defined as

$$\tilde{\boldsymbol{\xi}}_x \triangleq \hat{\boldsymbol{\xi}}_x - \boldsymbol{\xi}_x \tag{11a}$$

$$\overline{\boldsymbol{\Sigma}}_{\tilde{\xi}_x\tilde{\xi}_x} \triangleq \lim_{N_T \to \infty} N_T \mathcal{E}\langle \tilde{\boldsymbol{\xi}}_x \tilde{\boldsymbol{\xi}}_x^{\mathsf{H}} \rangle. \tag{11b}$$

Recall the vectorization of sample mean covariances

$$\hat{\boldsymbol{\xi}}_x = \frac{1}{N_T} \sum_{n_T=1}^{N_T} \boldsymbol{v}_{\mathsf{c}} \left(\boldsymbol{x}[n_T] \boldsymbol{x}^{\mathsf{H}}[n_T] \right). \tag{12}$$

Substituting the above into $\boldsymbol{v}_{\mathsf{c}}(\mathbf{ABC}) = (\mathbf{C}^{\mathsf{T}} \otimes \mathbf{A}) \boldsymbol{v}_{\mathsf{c}}(\mathbf{B})$, its concatenation can be represented into two forms

$$\begin{aligned} \boldsymbol{v}_{\mathsf{c}} \left(\boldsymbol{x}[n_T] \boldsymbol{x}^{\mathsf{H}}[n_T] \right) &= (\boldsymbol{I} \otimes \boldsymbol{x}[n_T]) \boldsymbol{x}^*[n_T] \\ &= (\boldsymbol{x}^*[n_T] \otimes \boldsymbol{I}) \boldsymbol{x}[n_T]. \end{aligned} \tag{13}$$

Here we define $\boldsymbol{\Sigma}_{\hat{\xi}_x\hat{\xi}_x} \triangleq \mathcal{E}\langle \hat{\boldsymbol{\xi}}_x \hat{\boldsymbol{\xi}}_x^{\mathsf{H}} \rangle \in \mathbb{C}_{\mathbb{H}}^{N_E^2 \times N_E^2}$ so that

$$\boldsymbol{\Sigma}_{\hat{\xi}_x\hat{\xi}_x} = \frac{1}{N_T^2} \sum_{n_T=1}^{N_T} \sum_{\acute{n}_T=1}^{N_T} \mathcal{E}\langle \underbrace{(\boldsymbol{x}^*[n_T] \otimes \boldsymbol{I})}_{\mathbf{A}} \underbrace{\boldsymbol{x}[n_T]}_{\mathbf{B}} \underbrace{\boldsymbol{x}^{\mathsf{T}}[\acute{n}_T]}_{\mathbf{C}} \underbrace{(\boldsymbol{I} \otimes \boldsymbol{x}^{\mathsf{H}}[\acute{n}_T])}_{\mathbf{D}} \rangle. \tag{14}$$

Proceeding on the formula of factorizing the matrix-valued product of Gaussian random variables [59]

$$\begin{aligned} \mathcal{E}\langle \mathbf{ABCD} \rangle = {} & \mathcal{E}\langle \mathbf{AB} \rangle \mathcal{E}\langle \mathbf{CD} \rangle \\ & + \mathcal{E}\langle \mathbf{C} \otimes \mathbf{A} \rangle \mathcal{E}\langle \mathbf{D} \otimes \mathbf{B} \rangle \\ & + \mathcal{E}\langle \mathbf{A} \mathcal{E}\langle \mathbf{BC} \rangle \mathbf{D} \rangle \\ & - 2\mathcal{E}\langle \mathbf{A} \rangle \mathcal{E}\langle \mathbf{B} \rangle \mathcal{E}\langle \mathbf{C} \rangle \mathcal{E}\langle \mathbf{D} \rangle. \end{aligned} \tag{15}$$

one can achieve the followings directly

$$
\begin{aligned}
\mathcal{E}\langle \mathbf{AB}\rangle \mathcal{E}\langle \mathbf{CD}\rangle &\overset{\hat{\xi}_x\hat{\xi}_x}{=\!=} \mathcal{E}\langle(\boldsymbol{x}^*[n_T]\otimes \boldsymbol{I})\boldsymbol{x}[n_T]\rangle\, \mathcal{E}\left\langle \boldsymbol{x}^{\mathsf{T}}[\acute{n}_T](\boldsymbol{I}\otimes \boldsymbol{x}^{\mathsf{H}}[\acute{n}_T])\right\rangle \\
&= \mathcal{E}\left\langle \boldsymbol{\upsilon}_{\mathsf{c}}\left(\boldsymbol{x}[n_T]\boldsymbol{x}^{\mathsf{H}}[n_T]\right)\right\rangle \mathcal{E}\left\langle \boldsymbol{\upsilon}_{\mathsf{c}}^{\mathsf{H}}\left(\boldsymbol{x}[\acute{n}_T]\boldsymbol{x}^{\mathsf{H}}[\acute{n}_T]\right)\right\rangle \\
&= \boldsymbol{\upsilon}_{\mathsf{c}}(\boldsymbol{\Sigma}_{xx})\,\boldsymbol{\upsilon}_{\mathsf{c}}^{\mathsf{H}}(\boldsymbol{\Sigma}_{xx}) \\
&= \boldsymbol{\xi}_x\boldsymbol{\xi}_x^{\mathsf{H}} \qquad (16a)\\
\mathcal{E}\langle \mathbf{C}\otimes\mathbf{A}\rangle \mathcal{E}\langle \mathbf{D}\otimes\mathbf{B}\rangle &\overset{\hat{\xi}_x\hat{\xi}_x}{=\!=} \mathcal{E}\left\langle \boldsymbol{x}^{\mathsf{T}}[\acute{n}_T]\otimes(\boldsymbol{x}^*[n_T]\otimes\boldsymbol{I})\right\rangle \mathcal{E}\left\langle (\boldsymbol{I}\otimes\boldsymbol{x}^{\mathsf{H}}[\acute{n}_T])\otimes\boldsymbol{x}[n_T]\right\rangle \\
&= \mathcal{E}\left\langle (\boldsymbol{x}^*[n_T]\boldsymbol{x}^{\mathsf{T}}[\acute{n}_T])\otimes\boldsymbol{I}\right\rangle \mathcal{E}\left\langle \boldsymbol{I}\otimes(\boldsymbol{x}[n_T]\boldsymbol{x}^{\mathsf{H}}[\acute{n}_T])\right\rangle \\
&= (\delta_{n_T,\acute{n}_T}\boldsymbol{\Sigma}_{xx}^{\mathsf{T}}\otimes\boldsymbol{I})(\boldsymbol{I}\otimes\delta_{n_T,\acute{n}_T}\boldsymbol{\Sigma}_{xx}) \\
&= \delta_{n_T,\acute{n}_T}(\boldsymbol{\Sigma}_{xx}^{\mathsf{T}}\otimes\boldsymbol{\Sigma}_{xx}) \qquad (16b)\\
\mathcal{E}\langle \mathbf{A}\mathcal{E}\langle \mathbf{BC}\rangle \mathbf{D}\rangle &\overset{\hat{\xi}_x\hat{\xi}_x}{=\!=} \mathcal{E}\left\langle (\boldsymbol{x}^*[n_T]\otimes\boldsymbol{I})\mathcal{E}\left\langle \boldsymbol{x}[n_T]\boldsymbol{x}^{\mathsf{T}}[\acute{n}_T]\right\rangle(\boldsymbol{I}\otimes\boldsymbol{x}^{\mathsf{H}}[\acute{n}_T])\right\rangle \\
&= \delta_{n_T,\acute{n}_T}\mathcal{E}\left\langle (\boldsymbol{x}^*[n_T]\otimes\boldsymbol{I})\boldsymbol{\Gamma}_{xx}(\boldsymbol{I}\otimes\boldsymbol{x}^{\mathsf{H}}[\acute{n}_T])\right\rangle \\
&= \left(\boldsymbol{\upsilon}_{\mathsf{c}}(\boldsymbol{\Gamma}_{xx})\,\boldsymbol{\upsilon}_{\mathsf{c}}^{\mathsf{H}}(\boldsymbol{\Gamma}_{xx})\right)^{\mathsf{BT}} \\
&= \delta_{n_T,\acute{n}_T}(\boldsymbol{\gamma}_x\boldsymbol{\gamma}_x^{\mathsf{H}})^{\mathsf{BT}} \qquad (16c)
\end{aligned}
$$

where $(\cdot)^{\mathsf{BT}}$ stands for block transposition, $\boldsymbol{\Gamma}_{xx} \triangleq \mathcal{E}\left\langle \boldsymbol{x}[n_T]\boldsymbol{x}^{\mathsf{T}}[n_T]\right\rangle \in \mathbb{C}_{\mathbb{S}}^{N_E\times N_E}$ and $\boldsymbol{\gamma}_x \triangleq \mathcal{E}\langle \boldsymbol{x}^*[n_T]\otimes\boldsymbol{x}[n_T]\rangle \in \mathbb{C}^{N_E^2\times 1}$ are the complementary covariance matrix and vector. Since Gaussian vector $\boldsymbol{x}[n_T]$ are of circularly symmetric complex-valued Gaussian random variable with zero-mean *i.e.* $\boldsymbol{\Gamma}_{xx} = \boldsymbol{O}$ or equivalently $\boldsymbol{\gamma}_x = \boldsymbol{0}$, such results would allow us to

$$
\begin{aligned}
\boldsymbol{\Sigma}_{\hat{\xi}_x\hat{\xi}_x} &= \frac{1}{N_T^2}\sum_{n_T=1}^{N_T}\sum_{\acute{n}_T=1}^{N_T}\boldsymbol{\xi}_x\boldsymbol{\xi}_x^{\mathsf{H}} + \delta_{n_T,\acute{n}_T}(\boldsymbol{\Sigma}_{xx}^{\mathsf{T}}\otimes\boldsymbol{\Sigma}_{xx}) \\
&= \boldsymbol{\xi}_x\boldsymbol{\xi}_x^{\mathsf{H}} + \frac{1}{N_T^2}\sum_{n_T=1}^{N_T}\boldsymbol{\Sigma}_{xx}^{\mathsf{T}}\otimes\boldsymbol{\Sigma}_{xx} \\
&= \boldsymbol{\xi}_x\boldsymbol{\xi}_x^{\mathsf{H}} + \frac{1}{N_T}\boldsymbol{\Psi}_{xx}.
\end{aligned} \qquad (17)
$$

The covariance matrices due to sample mean errors, $\boldsymbol{\Sigma}_{\tilde{\xi}_x\tilde{\xi}_x} = \mathcal{E}\left\langle(\hat{\boldsymbol{\xi}}_x - \boldsymbol{\xi}_x(\boldsymbol{\vartheta}_\circ))(\hat{\boldsymbol{\xi}}_x - \boldsymbol{\xi}_x(\boldsymbol{\vartheta}_\circ))^{\mathsf{H}}\right\rangle = \boldsymbol{\Sigma}_{\hat{\xi}_x\hat{\xi}_x} - \boldsymbol{\xi}_x\boldsymbol{\xi}_x^{\mathsf{H}}$, therefore, results in

$$
\overline{\boldsymbol{\Sigma}}_{\tilde{\xi}_x\tilde{\xi}_x} = \boldsymbol{\Psi}_{xx}. \qquad (18)
$$

Plugging (18) into (10), we arrive at

$$
\left[\overline{\boldsymbol{Q}}_{\mathrm{AML}}(\boldsymbol{\theta})\right]_{[n,\acute{n}]} = \dot{\boldsymbol{\xi}}_x^{\mathsf{H}}(\theta_n)\,\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\,\boldsymbol{\Pi}^{\perp}_{\mathcal{S}_{\mathcal{R}}\left(\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\boldsymbol{\Omega}\right)}\,\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}\dot{\boldsymbol{\xi}}_x(\theta_{\acute{n}}). \qquad (19)
$$

Since $\left[\bar{\boldsymbol{H}}_{\mathrm{AML}}(\boldsymbol{\theta})\right]_{[n,\acute{n}]}$ and $\left[\overline{\boldsymbol{Q}}_{\mathrm{AML}}(\boldsymbol{\theta})\right]_{[n,\acute{n}]}$ are the same for all n and $\acute{n}$, the quotation $\bar{\boldsymbol{H}}_{\mathrm{AML}}(\boldsymbol{\theta}) = \overline{\boldsymbol{Q}}_{\mathrm{AML}}(\boldsymbol{\theta})$ is verifiable. Along with (3.41) and (3.19), we obtain (3.42).

Proof of Theorem 2

Notice that $\boldsymbol{\Sigma}_{xx}$ might be rewritten as

$$\boldsymbol{\Sigma}_{xx} = \sum_{n_S=1}^{N_S} |s_{n_S}[n_T]|^2 \boldsymbol{\Sigma}_{hh}(\rho_{n_S}, \omega_{n_S}, \sigma_{\omega_{n_S}}) + \sigma_n^2 \boldsymbol{I}. \tag{20}$$

Then, to proof the Toeplitz structure in $\boldsymbol{\Sigma}_{xx}$, it suffices to proof that $\boldsymbol{\Sigma}_{hh}$ is Toeplitz. As seen in (2.6), the $(n_E, \acute{n}_E)$-th element of $\boldsymbol{\Sigma}_{hh}$ can be expressed as

$$\begin{aligned}
[\boldsymbol{\Sigma}_{hh}(\rho, \phi, \sigma_\phi)]_{[n_E, \acute{n}_E]} &= \rho \int f(\delta_\phi | 0; \sigma_\phi^2) e^{\imath k (n_E - 1) d_E \sin(\phi + \delta_\phi)} e^{-\imath k (\acute{n}_E - 1) d_E \sin(\phi + \delta_\phi)} d\delta_\phi \\
&= \rho \int f(\delta_\phi | 0; \sigma_\phi^2) e^{\imath k (n_E - \acute{n}_E) d_E \sin(\phi + \delta_\phi)} d\delta_\phi \\
&= \rho \int f(\delta_\phi | 0; \sigma_\phi^2) e^{\imath k ((n_L + 1) - 1) d_E \sin(\phi + \delta_\phi)} d\delta_\phi \\
&= [\boldsymbol{\Sigma}_{hh}(\rho, \phi, \sigma_\phi)]_{[n_L + 1, 1]}
\end{aligned} \tag{21}$$

which enables the channel covariance matrix $\boldsymbol{\Sigma}_{hh}$ to be Toeplitz as well.

Proof of Lemma 3

The solution (3.45) is easily derived by replacing $\boldsymbol{W}$ in (2.52) with $\boldsymbol{I}$. To verify (3.46), we reformulate the residual norm as

$$\begin{aligned}
\hat{\boldsymbol{\tau}}_{\mathsf{CM}} &= \arg\min_{\boldsymbol{\tau}} \| \boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Sigma}}_{xx} \|_{\mathsf{F}}^2 \\
&= \arg\min_{\boldsymbol{\tau}} \lceil (\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Sigma}}_{xx})^2 \rfloor.
\end{aligned} \tag{22}$$

where $\|\mathbf{A}\|_{\mathsf{F}}^2 \triangleq \lceil \mathbf{A}^{\mathsf{H}} \mathbf{A} \rfloor$ and $\lceil \mathbf{A} \rfloor$ denote the Frobenius norm and matrix trace respectively. Since the analytic function of CM needs to be satisfied by $\frac{\partial}{\partial \tau_{n_L}^*} \lceil (\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Sigma}}_{xx})^2 \rfloor = 0 \; ; \forall n_L$ [60, p. 891], its complex-valued derivative with respect to the n_L-th Toeplitz lag results in

$$\frac{\partial}{\partial \tau_{n_L}^*} \lceil (\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Sigma}}_{xx})^2 \rfloor = -2 \lceil \hat{\boldsymbol{\Sigma}}_{xx} \dot{\boldsymbol{\Sigma}}_{xx}(\tau_{n_L}^*) \rfloor + 2 \lceil \boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) \dot{\boldsymbol{\Sigma}}_{xx}(\tau_{n_L}^*) \rfloor \tag{23}$$

where $\dot{\mathbf{A}}(\mathrm{x}) \triangleq \frac{\partial}{\partial \mathrm{x}} \mathbf{A}(\mathrm{x})$ is the derivative with respect to x. Let us represent $\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau})$ in a linear structure according to

$$\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) = \tau_0 \boldsymbol{I} + \sum_{n_L=1}^{N_E - 1} \tau_{n_L} \boldsymbol{L}_{n_L} + \tau_{n_L}^* \boldsymbol{L}_{n_L}^{\mathsf{T}}. \tag{24}$$

Then, it follows that

$$\dot{\boldsymbol{\Sigma}}_{xx}(\tau^*_{n_L}) = \boldsymbol{L}^{\mathsf{T}}_{n_L}. \tag{25}$$

Forcing $\frac{\partial}{\partial \tau^*_{n_L}} \lceil (\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Sigma}}_{xx})^2 \rfloor \trianglerighteq 0$, we obtain the critical condition of the n_L-th Toeplitz lag as

$$\lceil \boldsymbol{L}_{n_L} \hat{\boldsymbol{\Sigma}}_{xx} \rfloor = \lceil \boldsymbol{L}_{n_L} \boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) \rfloor. \tag{26}$$

Proceeding on the n_L-th subdiagonal, it yields

$$\hat{\tau}_{n_L} = \frac{1}{N_E - n_L} \sum_{n_l=1}^{N_E - n_L} [\hat{\boldsymbol{\Sigma}}_{xx}]_{[n_L + n_l, n_l]} \tag{27}$$

which coincides with one available from the RA lag in (2.49).

Proof of Lemma 4

Inserting (2.48) into (3.51), the WLS function is

$$f_{\mathrm{WLS}}(\boldsymbol{\tau}|\hat{\boldsymbol{W}}) = \tilde{\boldsymbol{\xi}}_x^{\mathsf{H}} \hat{\boldsymbol{\Psi}}_{xx}^{-1} \tilde{\boldsymbol{\xi}}_x \tag{28}$$

where $\tilde{\boldsymbol{\xi}}_x \triangleq \hat{\boldsymbol{\xi}}_x - \breve{\boldsymbol{\Xi}} \boldsymbol{\Upsilon} \boldsymbol{\tau}$. Invoking the chain rule of $\frac{\partial}{\partial \boldsymbol{\tau}^{\mathsf{T}}} f_{\mathrm{WLS}}(\boldsymbol{\tau}|\hat{\boldsymbol{W}}) = \frac{\partial}{\partial \tilde{\boldsymbol{\xi}}_x^{\mathsf{H}}} f_{\mathrm{WLS}}(\boldsymbol{\tau}|\hat{\boldsymbol{W}}) \frac{\partial}{\partial \boldsymbol{\tau}^{\mathsf{T}}} \tilde{\boldsymbol{\xi}}_x$, it results in

$$\frac{\partial}{\partial \boldsymbol{\tau}^{\mathsf{T}}} f_{\mathrm{WLS}}(\boldsymbol{\tau}|\hat{\boldsymbol{W}}) = (2 \tilde{\boldsymbol{\xi}}_x^{\mathsf{H}} \hat{\boldsymbol{\Psi}}_{xx}^{-1})(-\breve{\boldsymbol{\Xi}} \boldsymbol{\Upsilon}). \tag{29}$$

Forcing $\frac{\partial}{\partial \boldsymbol{\tau}} f_{\mathrm{WLS}}(\boldsymbol{\tau}|\hat{\boldsymbol{W}}) \trianglerighteq \boldsymbol{0}$, we obtain

$$\hat{\boldsymbol{\tau}}_{\mathrm{WCM}} = (\boldsymbol{\Upsilon}^{\mathsf{H}} \breve{\boldsymbol{\Xi}}^{\mathsf{T}} \hat{\boldsymbol{\Psi}}_{xx}^{-1} \breve{\boldsymbol{\Xi}} \boldsymbol{\Upsilon})^{-1} \boldsymbol{\Upsilon}^{\mathsf{H}} \breve{\boldsymbol{\Xi}}^{\mathsf{T}} \hat{\boldsymbol{\Psi}}_{xx}^{-1} \hat{\boldsymbol{\xi}}_x \tag{30}$$

which coincides, through (2.48), with (3.50).

Coherent Source Localization via a Spatial Smoothing with Temporal Correlation

Bamrung Täu Sieskul Jatupon Pattanavichate Somchai Jitapunkul
Digital Signal Processing Research Laboratory, Department of Electrical Engineering, Chulalongkorn University
Bangkok, Thailand, e-mail: bts@chula.com, jatupal@chula.com, somchai.j@chula.ac.th

Abstract— **This paper aims at presenting a modified spatial smoothing for direction finding in the presence of coherent signals. Motivation of the approach stems from an incorporation of temporal correlation inherent in observable snapshots and simple quadratic covariance matrix. Under an appropriate arrangement of both, not only the temporal crosscorrelations are taken into account to refine the most appropriate rank of the array covariance matrix, but also the squared covariance matrix is constituted for enhancement of the angular resolution. In conjunction with root-MUSIC algorithm, the numerical results is performed to validate the significant improvement on DOA estimates.**

I. Introduction

Sensor array processing plays a prominent role in the phenomenons of the propagation of plane waves transmitting through a homogenous media and then impinging on an array antenna or sensor array. The problem of finding their directions, called *direction finding*, is of interest long times ago since it is a useful parameter in several systems such as wireless commination, radar, navigation and *etc.* In the presence of severely correlated signals, most of usual subspace-based estimators does not operate well, *i.e.* as same as in uncorrelated case because the array covariance matrix is *rank-deficient*. Various considerable ways for solving such a problem are to decorrelate the signals by eliminating the spatial crosscorrelations of correlated signals. Even in the coherent scenarios, it is conceivable that the spatial smoothing (SS) technique which makes use of averaging all subarray covariance matrices usually completes the rank of array covariance matrix.

Most related works involved spatial smoothing are, in general, to increase the DOA resolution (see *e.g.*[1]-[3]) in coherent case. When the signals are temporally correlated, the use of exploration on such a correlation should be more suitable than another one missing this additional information. Particularly, the more number of temporal lags has to retrieve more the diminish rank due to spatially correlated signals. As investigated in [4], it is evident that invoking the temporal crosscorrelations was able to more decorrelate the coherent signals. In addition, the DOA resolution by a quadratic spatial smoothing [3] could also be increased considerably because squaring the array covariance matrix yields a better condition from that ill-conditioned matrix due to the signal coherency.

Here we propose a combination between exploration of temporal correlation and quadratic covariance matrix under a variational procedure which does not affect to each beneficial performance. To see this, we first formulate the temporal covariance lags according to [4]. Individual exploitations of a spatial smoothing technique to all temporal lag matrices are disregarded, since it is not enough to increase the rank of final cross-temporal covariance matrix. On the other hand, we suggest to perform cross-temporal covariance approximation before spatial smoothing since it has to retrieve more the absent rank. Intuitively, since the proposed method requires only one time for spatial smoothing, the computational cost utilized by our algorithm seems to be less than that required by the previous approach [4].

The remainder of this paper begins with section II which introduces the model of passive sensor array that employs the uniform *l*inear array (ULA) receiving the discrete point source signals. In section III, the temporal correlations are first explored and then combined in the sense of crosscorrelation as a new sample covariance. Certainly, in section IV the quadratic spatial smoothing based on exploration of temporal correlations is implicitly proposed. To validate the so-called *f*orward-*e*xchange *q*udratic *a*uto *s*patial *s*moothing with *t*emporal *c*orrelation (TC-FEQASS) of our proposition, some numerical examples demonstrate improvements over the absence of exploiting the temporal correlations and/or the quadratic multiplication in section V. Finally, the summarization at section VI is given for the improved spatial smoothing.

II. Passive Array Model for Discrete Sources

Restrict ourselves to the propagation of $N_S \in \mathbb{N}^{1\times 1}$ complex wavefronts $s_{n_S}(t)e^{j(2\pi f_c t+\varphi_{n_S})}; n_S \in \{1,\ldots,N_S\}$ whose central frequency in narrowband assumption and initial phase carrier are $f_c \in \mathbb{R}_+^{1\times 1}$ and $\varphi_{n_S} \in [-\pi,\pi]$, respectively. Based on the superposition theorem, these wavefields all impinge on a ULA composing of $N_E \in \mathbb{N}^{1\times 1}$ omnidirectional sensor elements with an equidistance $d \in \mathbb{R}_+^{1\times 1}$. The sensor response is, in general, characterized by a steering vector $\boldsymbol{a}(\phi) : \mathbb{R}^{1\times 1} \mapsto \mathbb{C}^{N_E\times 1}$ at the *f*irst *e*lement *r*eference (FER)

$$\boldsymbol{a}(\phi) \triangleq \begin{bmatrix} 1 & e^{-jkd\sin(\phi)} & \ldots & e^{-jk(N_E-1)d\sin(\phi)} \end{bmatrix}^{\mathrm{T}} \quad (1)$$

with the wave number $k = 2\pi/\lambda$, the wavelength $\lambda = c/f_c$ and the light speed $c \simeq 3\times 10^8$. Collecting N_E measures under a fixed relation $t = n_T T_s$ of the sample period $T_s \in \mathbb{R}_+^{1\times 1}$ and the sample index $n_T \in \{n \,|\, 1 \le n \le N_T, n \in \mathbb{N}^{1\times 1}\}$ in $N_T \in \mathbb{N}^{1\times 1}$ samples, the array observation $\boldsymbol{x}[n_T] \in \mathbb{C}^{N_E\times 1}$ annoyed by a noise $\boldsymbol{n}[n_T] \in \mathbb{C}^{N_E\times 1}$ becomes

$$\boldsymbol{x}[n_T] = \boldsymbol{A}(\boldsymbol{\phi})\boldsymbol{s}[n_T] + \boldsymbol{n}[n_T] \quad (2)$$

where $\boldsymbol{A}(\boldsymbol{\phi}) : \mathbb{R}^{N_S\times 1} \mapsto \mathbb{C}_{\mathbb{V}}^{N_E\times N_S}$ signifies the array response matrix which is assumably time-invariant to all arrival directions $\phi_{n_S} \in [-\pi, \pi]$. Occasionally, the noise $\boldsymbol{n}[n_T]$ is modelled to be identically distributed as circularly symmetric Gaussian with zero mean and variance $\sigma_{\mathrm{n}}^2 \in \mathbb{R}_+^{1\times 1}$, *i.e.* $\boldsymbol{n}[n_T] \backsim \mathcal{CN}_{sc}(\boldsymbol{0}, \sigma_{\mathrm{n}}^2 \boldsymbol{I})$, so that it is also uncorrelated with $\boldsymbol{s}[n_T]$:

$$\mathcal{E}\langle \boldsymbol{s}[n_T]\boldsymbol{n}^{\mathrm{T}}[\acute{n}_T]\rangle = \boldsymbol{O}, \mathcal{E}\langle \boldsymbol{s}[n_T]\boldsymbol{n}^{\mathrm{H}}[\acute{n}_T]\rangle = \boldsymbol{O}$$
$$\mathcal{E}\langle \boldsymbol{n}[n_T]\boldsymbol{n}^{\mathrm{T}}[\acute{n}_T]\rangle = \boldsymbol{O}, \mathcal{E}\langle \boldsymbol{n}[n_T]\boldsymbol{n}^{\mathrm{H}}[\acute{n}_T]\rangle = \delta_{n_T,\acute{n}_T}\sigma_{\mathrm{n}}^2\boldsymbol{I}$$

where $\mathcal{E}\langle\cdot\rangle$ and $\delta_{n_T,\acute{n}_T}$ signify the statistical expectation and the Kronecker delta function, respectively. The array covariance matrix is referred to deal with the positive semi-definite Toeplitz Hermitian covariance matrix $\boldsymbol{R}_{xx} \in \mathbb{C}_{\mathrm{PTH}}^{N_E \times N_E}$

$$\boldsymbol{R}_{xx} \triangleq \mathcal{E}\langle \boldsymbol{x}[n_T]\boldsymbol{x}^{\mathrm{H}}[n_T]\rangle = \boldsymbol{A}(\boldsymbol{\phi})\boldsymbol{R}_{ss}\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{\phi}) + \sigma_{\mathrm{n}}^2\boldsymbol{I} \quad (3)$$

where $\boldsymbol{R}_{ss} \triangleq \mathcal{E}\langle \boldsymbol{s}[n_T]\boldsymbol{s}^{\mathrm{H}}[n_T]\rangle \in \mathbb{C}_{\mathrm{H}}^{N_S \times N_S}$ is the source signal covariance matrix. In fact, the available measure from the array might be tapped within a batch sample $\boldsymbol{X} \in \mathbb{C}^{N_E \times N_T}$

$$\boldsymbol{X} \triangleq \begin{bmatrix} \boldsymbol{x}[1] & \boldsymbol{x}[2] & \cdots & \boldsymbol{x}[N_T] \end{bmatrix} \quad (4)$$

Let $\hat{\boldsymbol{R}}_{xx} \in \mathbb{C}_{\mathrm{H}}^{N_E \times N_E}$ be the sample mean estimate of the array covariance matrix, defined by

$$\hat{\boldsymbol{R}}_{xx} \triangleq \frac{1}{N_T}\sum_{n_T=1}^{N_T} \boldsymbol{x}[n_T]\boldsymbol{x}^{\mathrm{H}}[n_T] = \frac{1}{N_T}\boldsymbol{X}\boldsymbol{X}^{\mathrm{H}} \quad (5)$$

To explore the implicit subspaces in raw data, we would decompose the covariance matrix $\hat{\boldsymbol{R}}_{xx}$ into a canonical form. If the eigenvalues are necessary for estimating the number of signals $\hat{N}_S$, we may adopt the OEVD providing $\hat{\boldsymbol{E}}_{\mathrm{s}} \in \mathbb{C}^{N_E \times \hat{N}_S}$ and $\hat{\boldsymbol{E}}_{\mathrm{n}} \in \mathbb{C}^{N_E \times (N_E - \hat{N}_S)}$ whose columns span the signal and noise subspace, respectively, resided in $\mathbb{C}^{N_E}$ (see *e.g.* [6]). Both of them have to be, in principle, aligned in

$$\begin{aligned}\hat{\boldsymbol{R}}_{xx} &\overset{\text{OEVD}}{=\!=} \hat{\boldsymbol{E}}\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{E}}^{\mathrm{H}} = \left[\hat{\boldsymbol{E}}_{\mathrm{s}} \vdots \hat{\boldsymbol{E}}_{\mathrm{n}}\right]\begin{bmatrix}\hat{\boldsymbol{\Lambda}}_{\mathrm{s}} & \boldsymbol{O}\\ \boldsymbol{O} & \hat{\boldsymbol{\Lambda}}_{\mathrm{n}}\end{bmatrix}\left[\hat{\boldsymbol{E}}_{\mathrm{s}} \vdots \hat{\boldsymbol{E}}_{\mathrm{n}}\right]^{\mathrm{H}} \\ &= \hat{\boldsymbol{E}}_{\mathrm{s}}\hat{\boldsymbol{\Lambda}}_{\mathrm{s}}\hat{\boldsymbol{E}}_{\mathrm{s}}^{\mathrm{H}} + \hat{\boldsymbol{E}}_{\mathrm{n}}\hat{\boldsymbol{\Lambda}}_{\mathrm{n}}\hat{\boldsymbol{E}}_{\mathrm{n}}^{\mathrm{H}}\end{aligned} \quad (6)$$

where $\hat{\boldsymbol{\Lambda}}_{\mathrm{s}} \in \mathbb{R}_{\mathbb{D}}^{\hat{N}_S \times \hat{N}_S}$ is the signal eigenvalue matrix whose diagonal contains the $\hat{N}_S$ largest eigenvalues $\{\hat{\lambda}_{n_E}\}_{n_E=1}^{\hat{N}_S}$ of $\hat{\boldsymbol{R}}_{xx}$ and $\hat{\boldsymbol{\Lambda}}_{\mathrm{n}} \in \mathbb{R}_{\mathbb{D}}^{(N_E - \hat{N}_S)\times(N_E - \hat{N}_S)}$ is the noise eigenvalue matrix whose diagonal contains the $(N_E - \hat{N}_S)$ smallest eigenvalues $\{\hat{\lambda}_{n_E}\}_{n_E=\hat{N}_S+1}^{N_E}$ of $\hat{\boldsymbol{R}}_{xx}$. The estimate $\hat{N}_S$ of signals may be available from testing the hypothesis of $h[N_S] : \mathbb{N}^{1\times 1} \mapsto \mathbb{R}^{1\times 1}$

$$\hat{N}_S = \arg\min_{N_S \in \mathcal{D}_{N_S}} h[N_S] \,; \mathcal{D}_{N_S} \triangleq \{0, 1, \ldots, N_E - 1\} \quad (7)$$

For instance, the minimum description length (MDL) [5] is

$$h_{\mathrm{MDL}}[N_S] \triangleq N_T(N_E - N_S)\log\left(\frac{\mu_{\mathrm{a}}[N_S]}{\mu_{\mathrm{g}}[N_S]}\right) + \frac{1}{2}N_S(2N_E - N_S)\log N_T$$

where the algebraic mean $\mu_{\mathrm{a}}[N_S] : \mathbb{N}^{1\times 1} \mapsto \mathbb{R}_+^{1\times 1}$ and the geometric mean $\mu_{\mathrm{g}}[N_S] : \mathbb{N}^{1\times 1} \mapsto \mathbb{R}_+^{1\times 1}$ are given by

$$\mu_{\mathrm{a}}[N_S] \triangleq \frac{1}{N_E - N_S}\sum_{N_S+1}^{N_E}\hat{\lambda}_{n_E}, \mu_{\mathrm{g}}[N_S] \triangleq \left(\prod_{N_S+1}^{N_E}\hat{\lambda}_{n_E}\right)^{\frac{1}{N_E - N_S}}$$

The estimate $\hat{\boldsymbol{\phi}} \in \mathbb{R}^{\hat{N}_S \times 1}$ of the signal directions can be estimated from a subspace-based algorithms. Most subspace-based algorithms to concern with is that of MUSIC [6]

$$\hat{\phi}_{\text{MUSIC}} \triangleq \arg\min_{\phi}\left\|\boldsymbol{a}^{\mathrm{H}}(\phi)\hat{\boldsymbol{E}}_{\mathrm{n}}\right\|_{\mathrm{E}}^2 \quad (8)$$

III. Exploration of Temporal Crosscorrelations

In buffering all N_T snapshots, we may rearrange them into $N_W \in \mathbb{N}^{1\times 1}$ windows whose size $N_t \in \mathbb{N}^{1\times 1}$ are such that

$$N_T \le N_W N_t \quad ; N_W \triangleq N_T - N_t + 1 \quad (9)$$

For a smaller sample batch, we may collect the received batches $\boldsymbol{X}_{N_t}[n_W] \in \mathbb{C}^{N_E \times N_t}$ and $\boldsymbol{X}_{N_W}[n_t] \in \mathbb{C}^{N_E \times N_W}$ as

$$\boldsymbol{X}_{N_t}[n_W] \triangleq \begin{bmatrix}\boldsymbol{x}[n_W] & \boldsymbol{x}[n_W+1] & \cdots & \boldsymbol{x}[n_W+N_t-1]\end{bmatrix} \quad (10a)$$
$$\boldsymbol{X}_{N_W}[n_t] \triangleq \begin{bmatrix}\boldsymbol{x}[n_t] & \boldsymbol{x}[n_t+1] & \cdots & \boldsymbol{x}[n_t+N_W-1]\end{bmatrix} \quad (10b)$$

Theoretically, the emitted signal correlation is, in general, assumed circularly symmetric such that $\mathcal{E}\langle \boldsymbol{s}[n_T]\boldsymbol{s}^{\mathrm{H}}[\acute{n}_T]\rangle = \delta_{n_T,\acute{n}_T}\boldsymbol{R}_{ss}$ and $\mathcal{E}\langle \boldsymbol{s}[n_T]\boldsymbol{s}^{\mathrm{T}}[\acute{n}_T]\rangle = \boldsymbol{O}$. However, the first condition is relaxed herein so that $\mathcal{E}\langle \boldsymbol{s}[n_T]\boldsymbol{s}^{\mathrm{H}}[\acute{n}_T]\rangle = \boldsymbol{R}_{ss}[n_l]$ which is, in general, assumed a non-zero matrix for any time lag $n_l \triangleq \acute{n}_T - n_T$. Consider a cross-temporal covariance matrix $\boldsymbol{R}_{XX} \in \mathbb{C}_{\mathrm{H}}^{N_t N_E \times N_t N_E}$ due to all partitioned subwindows

$$\begin{aligned}\boldsymbol{R}_{XX} &\triangleq \mathcal{E}\left\langle \boldsymbol{v}_{\mathrm{c}}(\boldsymbol{X}_{N_t}[n_W])\boldsymbol{v}_{\mathrm{c}}^{\mathrm{H}}(\boldsymbol{X}_{N_t}[n_W])\right\rangle \\ &= \begin{bmatrix}\boldsymbol{R}_{xx}[0] & \boldsymbol{R}_{xx}^{\mathrm{H}}[1] & \cdots & \boldsymbol{R}_{xx}^{\mathrm{H}}[N_t-1]\\ \boldsymbol{R}_{xx}[1] & \boldsymbol{R}_{xx}[0] & \cdots & \boldsymbol{R}_{xx}^{\mathrm{H}}[N_t-2]\\ \vdots & \vdots & \ddots & \vdots\\ \boldsymbol{R}_{xx}[N_t-1] & \boldsymbol{R}_{xx}[N_t-2] & \cdots & \boldsymbol{R}_{xx}[0]\end{bmatrix}\end{aligned} \quad (11)$$

where each entry $\boldsymbol{R}_{xx}[n_l] : \mathbb{N}^{1\times 1} \mapsto \mathbb{C}^{N_E \times N_E}$ is defined by

$$\begin{aligned}\boldsymbol{R}_{xx}[n_l] &\triangleq \mathcal{E}\langle \boldsymbol{x}[n_t]\boldsymbol{x}^{\mathrm{H}}[\acute{n}_t]\rangle \\ &= \boldsymbol{A}(\boldsymbol{\phi})\boldsymbol{R}_{ss}[n_l]\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{\phi}) + \delta[n_l]\sigma_{\mathrm{n}}^2\boldsymbol{I}\end{aligned} \quad (12)$$

where $\delta[n_l]$ designates the dirac delta function. Comparing (3) with (12), it can be pointed out that

$$\boldsymbol{R}_{xx}[0] = \boldsymbol{R}_{xx} = \boldsymbol{A}(\boldsymbol{\phi})\boldsymbol{R}_{ss}\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{\phi}) + \sigma_{\mathrm{n}}^2\boldsymbol{I} \quad (13)$$

In finite time series, the temporal covariance is available from a finite sample covariance matrix $\hat{\boldsymbol{R}}_{XX} \in \mathbb{C}_{\mathrm{H}}^{N_t N_E \times N_t N_E}$ as

$$\begin{aligned}\hat{\boldsymbol{R}}_{XX} &\triangleq \frac{1}{N_W}\sum_{n_W=1}^{N_W}\boldsymbol{v}_{\mathrm{c}}(\boldsymbol{X}_{N_t}[n_W])\boldsymbol{v}_{\mathrm{c}}^{\mathrm{H}}(\boldsymbol{X}_{N_t}[n_W]) \\ &= \begin{bmatrix}\hat{\boldsymbol{R}}_{xx}[1,1] & \hat{\boldsymbol{R}}_{xx}^{\mathrm{H}}[2,1] & \cdots & \hat{\boldsymbol{R}}_{xx}^{\mathrm{H}}[N_t,1]\\ \hat{\boldsymbol{R}}_{xx}[2,1] & \hat{\boldsymbol{R}}_{xx}[2,2] & \cdots & \hat{\boldsymbol{R}}_{xx}^{\mathrm{H}}[N_t,2]\\ \vdots & \vdots & \ddots & \vdots\\ \hat{\boldsymbol{R}}_{xx}[N_t,1] & \hat{\boldsymbol{R}}_{xx}[N_t,2] & \cdots & \hat{\boldsymbol{R}}_{xx}[N_t,N_t]\end{bmatrix}\end{aligned} \quad (14)$$

where the entry $\hat{\boldsymbol{R}}_{xx}[n_t, \acute{n}_t] \in \mathbb{C}^{N_E \times N_E}$ becomes

$$\begin{aligned}\hat{\boldsymbol{R}}_{xx}[n_t, \acute{n}_t] &\triangleq \frac{1}{N_W}\sum_{n_W=1}^{N_W}\boldsymbol{x}[n_t+n_W-1]\boldsymbol{x}^{\mathrm{H}}[\acute{n}_t+n_W-1] \\ &= \frac{1}{N_W}\boldsymbol{X}_{N_W}[n_t]\boldsymbol{X}_{N_W}^{\mathrm{H}}[\acute{n}_t]\end{aligned} \quad (15)$$

One way to find $\boldsymbol{R}_{xx}$ appropriately is to evaluate the average of $\hat{\boldsymbol{R}}_{xx[n_t,n_t]};\forall n_t$. In [4], it has been argued to make use of the crosscovariances $\hat{\boldsymbol{R}}_{xx}[n_t,\acute{n}_t];\forall n_t \neq \acute{n}_t$, *i.e.* in off-diagonals of $\hat{\boldsymbol{R}}_{XX}$. Let $\hat{\boldsymbol{R}}_{\mathrm{TC}} \in \mathbb{C}_{\mathbb{H}}^{N_E \times N_E}$ be the covariance estimate by

$$\begin{aligned}\hat{\boldsymbol{R}}_{\mathrm{TC}} &\triangleq \frac{1}{N_t}\sum_{n_t=1}^{N_t}\sum_{\acute{n}_t=1}^{N_t}\hat{\boldsymbol{R}}_{xx}[n_t,\acute{n}_t]\hat{\boldsymbol{R}}_{xx}^{-1}[\acute{n}_t,\acute{n}_t]\hat{\boldsymbol{R}}_{xx}[\acute{n}_t,n_t] \\ &\approx \frac{1}{N_t^2}\sum_{n_t=1}^{N_t}\sum_{\acute{n}_t=1}^{N_t}\hat{\boldsymbol{R}}_{xx}^{\mathrm{H}}[\acute{n}_t,n_t]\hat{\boldsymbol{R}}_{xx}[\acute{n}_t,n_t]\end{aligned} \tag{16}$$

It is important to emphasize that (16) can increase the rank of array covariance matrix beneficially when N_t is large as possible as N_T *i.e.* as N_W approaches 1. However, the larger the window length, the greater the computational cost required in temporal-based spatial smoothing.

IV. Forward-Exchange Quadratic Auto Spatial Smoothing

With respect to the first element, the index of subarray $n_A \in \{n \mid 1 \leq n \leq N_A, n \in \mathbb{N}^{1\times 1}\}$ is characterized by a binary selection matrix $\boldsymbol{\Xi}[n_A] : \mathbb{N}^{1\times 1} \mapsto \mathbb{B}^{N_E \times N_e}$

$$\boldsymbol{\Xi}[n_A] \triangleq \underbrace{\begin{bmatrix} \boldsymbol{O} \\ \boldsymbol{I} \\ \boldsymbol{O} \end{bmatrix}}_{N_e} \begin{matrix} \}\, n_A - 1 \\ \}\, N_e \\ \}\, N_E - N_e - n_A + 1 \end{matrix} \tag{17}$$

The forward quadratic spatial smoothing covariance matrix of the n_Ath subarray, $\boldsymbol{R}_{\mathrm{FQ}}[n_A] \in \mathbb{C}_{\mathbb{H}}^{N_e \times N_e}$, can be evaluated from

$$\boldsymbol{R}_{\mathrm{FQ}}[n_A] \triangleq \boldsymbol{\Xi}^{\mathrm{T}}[n_A]\boldsymbol{R}_{xx}^2\boldsymbol{\Xi}[n_A] \quad ;\boldsymbol{R}_{xx}^2 \triangleq \boldsymbol{R}_{xx}\boldsymbol{R}_{xx} \tag{18}$$

where $\boldsymbol{R}_{xx}^2 \in \mathbb{C}_{\mathbb{H}}^{N_E \times N_E}$ denote the quadratic covariance matrix. Smoothing the above subarray covariance in spatial domain, we then have $\widetilde{\boldsymbol{R}}_{\mathrm{F}} \in \mathbb{C}_{\mathbb{H}}^{N_e \times N_e}$, the forward-only auto spatial smoothing covariance matrix, by performing

$$\begin{aligned}\widetilde{\boldsymbol{R}}_{\mathrm{FQ}} &\triangleq \frac{1}{N_A}\sum_{n_A=1}^{N_A}\boldsymbol{R}_{\mathrm{FQ}}[n_A] = \boldsymbol{\Xi}^{\mathrm{T}}(\boldsymbol{I}_{(N_A)} \otimes \boldsymbol{R}_{xx}^2)\boldsymbol{\Xi} \quad ; \\ \boldsymbol{\Xi} &\triangleq \frac{1}{\sqrt{N_A}}\begin{bmatrix}\boldsymbol{\Xi}^{\mathrm{T}}[1] & \boldsymbol{\Xi}^{\mathrm{T}}[2] & \cdots & \boldsymbol{\Xi}^{\mathrm{T}}[N_A]\end{bmatrix}^{\mathrm{T}}\end{aligned} \tag{19}$$

where $\boldsymbol{\Xi} \in \mathbb{C}^{N_A N_E \times N_e}$ is subarray selection matrix and $\otimes$ designates the Kronecker product. Let $\boldsymbol{i}_{[n_E]} \in \mathbb{R}^{N_E \times 1}$ be the n_Eth column of the identity matrix $\boldsymbol{I}_{(N_E)}$. The exchange matrix, $\acute{\boldsymbol{I}} \triangleq [\,\boldsymbol{i}_{[N_E]} \;\; \boldsymbol{i}_{[N_E-1]} \;\; \cdots \;\; \boldsymbol{i}_{[1]}\,] \in \mathbb{B}^{N_E \times N_E}$, is one to concern with the exchange portion of $\widetilde{\boldsymbol{R}}_{\mathrm{FQ}}$ (see *e.g.*[1]), $\widetilde{\boldsymbol{R}}_{\mathrm{EQ}} \in \mathbb{C}_{\mathbb{H}}^{N_e \times N_e}$, from

$$\widetilde{\boldsymbol{R}}_{\mathrm{EQ}} = \acute{\boldsymbol{I}}\boldsymbol{\Xi}^{\mathrm{T}}(\boldsymbol{I} \otimes (\boldsymbol{R}_{xx}^2)^*)\boldsymbol{\Xi}\acute{\boldsymbol{I}} \tag{20}$$

If $\acute{\boldsymbol{\Xi}} \triangleq \boldsymbol{\Xi}\acute{\boldsymbol{I}} \in \mathbb{R}^{N_A N_E \times N_e}$, the covariance matrix of forward-exchange auto spatial smoothing, $\widetilde{\boldsymbol{R}}_{\mathrm{FE}} \in \mathbb{C}_{\mathbb{H}}^{N_e \times N_e}$, is then

$$\widetilde{\boldsymbol{R}}_{\mathrm{FEQ}} = \frac{1}{2}\left(\boldsymbol{\Xi}^{\mathrm{T}}(\boldsymbol{I} \otimes \boldsymbol{R}_{xx}^2)\boldsymbol{\Xi} + \acute{\boldsymbol{\Xi}}^{\mathrm{T}}(\boldsymbol{I} \otimes (\boldsymbol{R}_{xx}^2)^*)\acute{\boldsymbol{\Xi}}\right) \tag{21}$$

V. TC-FEQASS Algorithm

To clarify our algorithm, the proposed algorithm assume that the estimated number of sources, $\hat{N}_S$ in (7), is strictly coincide with the true value N_S. The TC-FEQASS procedure follows from:

1) Find the window length N_t from $N_t = N_T - N_W + 1$.
2) Rearrange $\boldsymbol{X}$ of (4) into N_t batches as $\boldsymbol{X}_{N_W}[n_t]$ in (10).
3) Formulate all lower triangular entries of $\hat{\boldsymbol{R}}_{\mathrm{T}}$ by (15).
4) Perform $\hat{\boldsymbol{R}}_{\mathrm{TC}}$ in (16).
5) Let $\hat{\widetilde{\boldsymbol{R}}}_{\text{TC-FEQASS}} \in \mathbb{C}_{\mathbb{H}}^{N_e \times N_e}$ be the result from substituting $\boldsymbol{R}_{xx}$ in (21) with the calculated $\hat{\boldsymbol{R}}_{\mathrm{TC}}$.
6) Compute the estimate $\hat{\boldsymbol{\phi}}$ by a subspace-based algorithm.

For terminological concise, we refer the approach presented in [1] to as the *f*orward-*e*xchange *a*uto *s*patial *s*moothing (FE-ASS) which is equivalent to *f*orward-*b*ackward *a*uto *s*patial *s*moothing (FBASS) in [2]. We designate the approach presented in [3] to as *f*orward-*e*xchange *q*uadratic *a*uto *s*patial *s*moothing (FEQASS). Indeed, the last one to compare with the paper's approach is the *f*orward-*e*xchange *a*uto *s*patial *s*moothing with *t*emporal *c*orrelation (TC-FEASS) in [4]. It should be noted that in [4], all lower triangular entries of $\hat{\boldsymbol{R}}_{\mathrm{T}}$ are separately performed by *f*orward-*e*xchange *a*uto *s*patial *s*moothing (FEASS) at its step 4). Then, they are all temporally smoothed to be the crosscovariance matrix at its step 5) according to (16). Therefore, when N_t is large, the computational cost for separately smoothing those lower triangular entries seems to be much more than that required by our proposed TC-FEQASS.

VI. Numerical Examples

To demonstrate the impact of the proposed method to subspace-based algorithm, we commonly employ the original signal $\boldsymbol{s}_{\mathrm{o}}[n_T] \backsim \mathcal{CN}(\boldsymbol{0},\sigma_{\mathrm{s}}^2\boldsymbol{I})$ and their correlation coefficient $\rho \in \mathbb{C}^{1\times 1}$ in $\boldsymbol{s}[n_T] = \boldsymbol{\Gamma}(\rho)\boldsymbol{s}_{\mathrm{o}}[n_T]$, where $\boldsymbol{\Gamma}(\rho) : \mathbb{C}^{1\times 1} \mapsto \mathbb{C}^{N_S \times N_S}$ denotes the correlating characterization. Let the initial phase of all signals be uniformly distributed as $\varphi_{n_S} \backsim \mathcal{U}[-\pi,\pi]$ so that the down-converted envelopes obey $\{s_{n_S}(t)e^{j\varphi_{n_S}}\}_{n_S=1}^{N_S}$ and SNR $\triangleq 10\log\left(\sigma_{\mathrm{s}}^2/\sigma_{\mathrm{n}}^2\right)$ signifies their strength. Under a moderate angular separation with two angulars of arrival (AOA), the FER-ULA is incorporated with the root-MUSIC [7] and the following parameters

AOA	SNR	N_S	N_T	N_W	N_E	N_e
−5, 5	20	2	2^5	2	6	4

$$\begin{bmatrix} s_1[n_T] \\ s_2[n_T] \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-|\rho|^2} \end{bmatrix}\begin{bmatrix} s_{\mathrm{o}_1}[n_T] \\ s_{\mathrm{o}_2}[n_T] \end{bmatrix}$$

The values shown above are taken into account throughout the paper, unless otherwise a variation on the parameter of interest will be specified individually. It allows us to receive a standard covariance matrix of two equi-power signals

$$\boldsymbol{R}_{ss} = \begin{bmatrix} \sigma_{\mathrm{s}}^2 & \sigma_{\mathrm{s}}^2\rho^* \\ \sigma_{\mathrm{s}}^2\rho & \sigma_{\mathrm{s}}^2 \end{bmatrix}$$

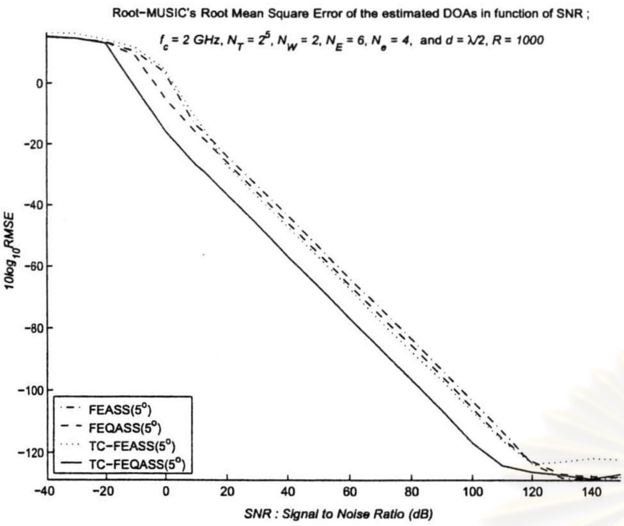


Fig. 1. SNR performance

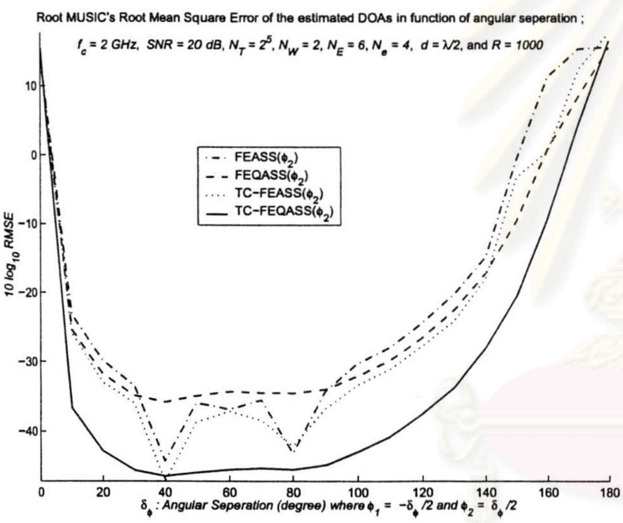


Fig. 2. Angular seperation performance

Owing to the results that both directional estimates are preliminarily observed to provide the root mean square error (RMSE) similarly, we then illustrate only one signal for each scenario. Investigating the SNR performance in figure 1, one can see that the TC-FEQASS yields the lowest SNR threshold and relative RMSE. Furthermore, the TC-FEASS is not superior to other none temporal-based spatial smoothing approaches at low SNR, especially less than 20 dB. In figure 2, the TC-FEQASS outperforms all other candidates without quadratic multiplication and temporal correlation. Lastly, putting two signals whose their directions are symmetric with respect to the array broadside in the experiment, the TC-FEQASS keeps the lowest RMSE obviously in figure 3. Monotonically, the larger the signal correlation amplitude, the greater the increase of each RMSE.

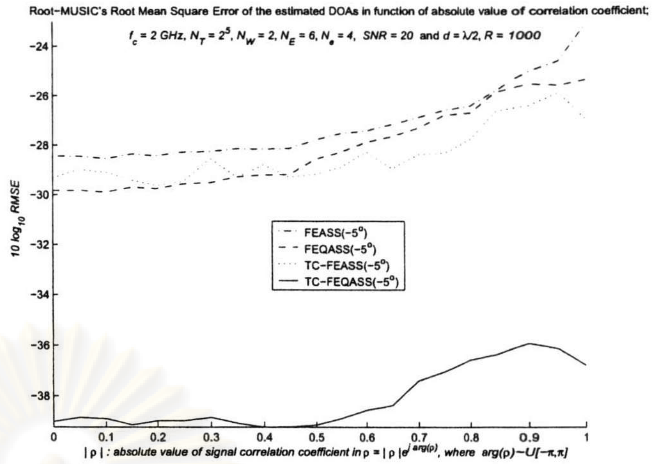


Fig. 3. Signal correlation performance

VII. Conclusion

The improvement of spatial smoothing based on temporal correlation and quadratic multiplication has been performed. The new approach is to rearrange the computations between forward-exchange spatial smoothing and cross-temporal correlations, and then make use of quadratic array covariance matrix by squaring. Unfortunately, the temporal-based spatial smoothing requires more computational cost. However, the rearranged computation shown above seems to be more efficient than that in [4] because there is only one time for spatial smoothing formulation. As seen numerically, the TC-FEQASS provides better performances for subspace-based DOA estimation.

Acknowledgment

The authors would like to express an appreciation to collaborators's encouragement and also thank to the financial support form the cooperative research unit between Chulalongkorn University and the Private Sector.

References

[1] R. T. Williams, S. Prasad, A. K. Mahalanabis, and L. H. Sibul, "An improved spatial smoothing technique for bearing estimation of coherent signals," *IEEE Transactions on Acoustic, Speech, and Signal Processing*, vol. ASSP-36, pp. 425-432, Apr. 1988.

[2] S. U. Pillai and B. H. Kwon, "Forward/backward spatial smoothing technique for coherent signal identification," *IEEE Transactions on Acoustic, Speech, and Signal Processing*, vol. ASSP-37, pp. 8-15, Jan. 1989.

[3] J. Li, "Improved angular resolution for spatial smoothing techniques," *IEEE Transactions on Signal Processing*, vol. SP-40, pp. 3078-3081, Dec. 1992.

[4] W. Du and R. L. Kirlin, "Enhancement of covariance matrix for array processing," *IEEE Transactions on Signal Processing*, vol. SP-40, pp. 2602-2606, Oct. 1992.

[5] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on Acoustic, Speech, and Signal Processing*, vol. ASSP-33, pp. 387-392, Apr. 1985.

[6] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *Proc. RADC Spectrum Estimation Workshop*, pp. 243-258, Oct. 1979.

[7] Barabell, A, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 8, pp. 336 -339 Apr. 1983.

Decoupled Estimation of Nominal Direction and Angular Spread based on Asymptotic Maximum Likelihood Approach

Bamrung Täu Sieskul
Center of Excellence in Telecommunication Technology
Chulalongkorn University
254 Phyathai Road, Patumwan, Bangkok
bamrung.t@student.chula.ac.th

Somchai Jitapunkul
Department of Electrical Engineering
Chulalongkorn University
254 Phyathai Road, Patumwan, Bangkok
somchai.j@chula.ac.th

Abstract: - The problem of estimating the nominal direction and its underlying angular spread is considered herein. As encountered in spatially distributed source localization, the computation of these directional parameters might be regarded as two consecutive tasks. In this paper, we propose an *a*symptotic *m*aximum *l*ikelihood (AML) approach to successively estimate both of them. The first advantage of estimation in this way is that it requires only two successive 1-dimensional searches rather than joint 2-dimensional optimization as utilized in the *a*symptotic *m*aximum *l*ikelihood (AML) estimator. Therefore, the decoupled estimation in this way provides more numerical flexibility. Since it belongs to a large-sample approximation of the exact ML method, numerical simulation is conducted in order to validate its asymptotic efficiency producible with respect to Cramér-Rao bound. Although its non-asymptotic performance is inferior to that provided by the joint AML approach, it appeared that, in the region of large number of temporal snapshot, the proposed AML estimator for decoupled estimation is the same as the AML criterion which employed the joint 2-dimensions.

Key-Words: - Maximum Likelihood, Parameter Estimation, Source Localization

1 Introduction

Most works involved direction finding problem were based on *m*aximum *l*ikelihood (ML) estimation due to its producible optimality [1]. To arrive at extremal quantity, an optimization search of the likelihood function seems to be inevitable in a complex model. In general, the physical model with well-described characterization would requires large number of model parameters. As a consequence, the larger the number of model parameters, the larger the dimension of optimization over parameter space. Unfortunately, this might allow the ML estimator to be unsuitable for being incorporated into real-word applications. This is because of suffering from implementation aspects, for instance, computational complexity and memory consumption.

In the presence of local scattering around the vicinity of source, most classical point source models will suffer from the lack of identifiability in the presence of large number of directions. To deviate from the given problem, it is reasonable to assume that a number of multipaths should be large enough so that their path gains can be characterized, under the central limit theorem, by a Gaussian random variable whose associated directions are also random [2]–[3]. With a priori knowledge of angle probability distribution, it appeared in general possible to govern all deviated angles into a parametric model as well. As a matter of course, an incoming source signal immediately consists itself of three individual arrival parameters, such as, nominal direction, angular spread and power observed by the sensor array. Since the exact likelihood function with all 4 parameters, which also include the spatially uncorrelated noise variance, can not be concentrated on explicitly [2], it is therefore conflictive to account for implementations as indicated before. Recently, a large-sample approximation of the exact ML is proposed in [4]. It requires joint 2-dimensional search and yields lower in error variances than the WLS (*w*eighted *l*east *s*quares) estimator in [2]. Furthermore, relying on these restrictions, the reducible computation might be admitted by two successive one-dimensional WLS searches [5] for estimating the nominal direction and angular spread.

Here we propose an *a*symptotic *m*aximum *l*ikelihood (AML) approach to successively estimate both of them. The first advantage of estimation in this way is that it requires only two successive 1-dimensional searches instead of joint 2-dimensional optimization as utilized in the AML estimator [4]. Since it belongs to a large-sample approximation of the exact ML method, numerical simulation is conducted in order to validate its asymptotic efficiency producible with respect to Cramér-Rao bound. Although its non-asymptotic performance is inferior to that provided by the AML approach, it appeared that the proposed AML estimator for decoupled estimation still keeps the asymptotic efficiency.

2 Spatially Distributed Source Model

Restrict our attention to a signal transmitting through a channel and then impinging on the *u*niform *l*inear *a*rray (ULA). With phase reference at the first element, or *f*irst element reference (FER), the array response vector $\boldsymbol{a}(\phi) : [-\frac{\pi}{2}, \frac{\pi}{2}] \mapsto \mathbb{C}^{N_E \times 1}$ can be, in general, written ideally as

$$\boldsymbol{a}(\phi) \triangleq \begin{bmatrix} 1 & e^{\imath k d_E \sin(\phi)} & \dots & e^{\imath k d_E (N_E - 1)\sin(\phi)} \end{bmatrix}^{\mathsf{T}} \quad (1)$$

where $k = \frac{2\pi}{\lambda}$ designates the wave number with associating wavelength λ and N_E is the number of sensor elements. As previously developed, most local scattering models assume that the nominal angle ϕ is deterministic while angular deviation δ_ϕ and associating path gain γ are considered as stochastic quantities. According to linear regression analysis, the array output at time instant n_T can be characterized in a flat fading channel by the snapshot $\boldsymbol{x}[n_T] \in \mathbb{C}^{N_E \times 1}$. Mathematically speaking, it can be represented as [3, p. 25]

$$\boldsymbol{x}[n_T] = s[n_T] \sum_{n_P=1}^{N_P} \gamma_{n_P}[n_T] \boldsymbol{a}(\phi + \delta_{\phi_{n_P}}[n_T]) + \boldsymbol{n}[n_T] \quad (2)$$

where N_P denotes the number of scattering paths and $\boldsymbol{n}[n_T] \in \mathbb{C}^{N_E \times 1}$ designates the additive noise at sensor array. For a large number of rays, the channel vector

$$\boldsymbol{h}[n_T] \triangleq \sum_{n_P=1}^{N_P} \gamma_{n_P}[n_T] \boldsymbol{a}(\phi + \delta_{\phi_{n_P}}[n_T]) \quad (3)$$

seemed, under the central limit theorem, plausible to hold a circularly-symmetric complex-valued Gaussian process, *i.e.*, $\boldsymbol{h}[n_T] \sim \mathcal{N}_c(\boldsymbol{0}; \boldsymbol{\Sigma}_{hh}, \boldsymbol{O})$. This N_E-dimensional variate implicitly provides the statistic $\boldsymbol{\Sigma}_{hh} \triangleq \mathcal{E}\langle \tilde{\boldsymbol{h}}[n_T]\tilde{\boldsymbol{h}}^{\mathsf{H}}[n_T] \rangle \in \mathbb{C}_{\mathbb{H}}^{N_E \times N_E}$, where $\tilde{\boldsymbol{h}}[n_T] \triangleq \boldsymbol{h}[n_T] - \mathcal{E}\langle \boldsymbol{h}[n_T] \rangle = \boldsymbol{h}[n_T]$. For taking an incoherently distributed channel into account, the second-order statistic of a certain incoming ray yields [3]

$$\mathcal{E}\langle \gamma_{n_P}[n_T] \gamma^*_{\acute{n}_P}[\acute{n}_T] \rangle = \sigma_\gamma^2 \delta_{n_P, \acute{n}_P} \delta_{n_T, \acute{n}_T} \quad (4)$$

where $\delta_{\bullet,\bullet}$ signifies the Krönecker delta function and σ_γ^2 is the power due to any path. Over spatial continuum of incoming rays, it can be approximated as

$$\boldsymbol{\Sigma}_{hh}(\rho, \phi, \sigma_\phi) \approx \rho \int f(\delta_\phi | 0; \sigma_\phi^2) \boldsymbol{a}(\phi + \delta_\phi) \boldsymbol{a}^{\mathsf{H}}(\phi + \delta_\phi) d\delta_\phi \quad (5)$$

where $\rho \triangleq N_P \sigma_\gamma^2$ signifies the cluster power due to all paths and $f(\delta_\phi | 0; \sigma_\phi^2)$ denotes the conditional PDF for random deviation δ_ϕ given a priori knowledge of the angular spread σ_ϕ. In instead of such physical angles ϕ and σ_ϕ, the spatial frequency response is preferable due to the better accuracy of approximating the first-order Taylor series around the array broadside [3]. In general, the spatial frequency ω and its associating standard deviation σ_ω are provided by

$$\omega(\phi) = k d_E \sin(\phi) \quad (6a)$$

$$\sigma_\omega(\phi, \sigma_\phi) = k d_E \cos(\phi) \sigma_\phi. \quad (6b)$$

Accounting for small angular spread, the so-called *spatial frequency* approximation results in a separable form as

$$\boldsymbol{\Sigma}_{hh}(\rho, \omega, \sigma_\omega) \simeq \rho \boldsymbol{D}_a(\omega) \boldsymbol{B}(\sigma_\omega) \boldsymbol{D}_a^{\mathsf{H}}(\omega) \quad (7)$$

where $\boldsymbol{D}_a(\omega) : [-kd_E, kd_E] \mapsto \mathbb{C}_{\mathbb{D},\mathbb{U}}^{N_E \times N_E}$ is diagonal and unitary matrix parameterized by nominal angle and $\boldsymbol{B}(\sigma_\omega) : \mathbb{R}_+^{1 \times 1} \mapsto \mathbb{R}_{\mathbb{S},\mathbb{T}}^{N_E \times N_E}$ is symmetric Toeplitz matrix parameterized by angular spread. Their $(n_E, \acute{n}_E)$-th elements can be expressed by [3, p. 22])

$$[\boldsymbol{D}_a(\omega)]_{[n_E, \acute{n}_E]} = e^{\imath (n_E - 1)\omega} \delta_{n_E, \acute{n}_E} \quad (8a)$$

$$[\boldsymbol{B}(\sigma_\omega)]_{[n_E, \acute{n}_E]} = f_{\mathcal{F}}((n_E - \acute{n}_E)\sigma_\omega | 0, 1) \quad (8b)$$

whence characteristic function $f_{\mathcal{F}}(t|, 0, 1) \triangleq \mathcal{F}(f(\delta_\omega|, 0, 1))$ is equivalent to the Fourier transform $\mathcal{F}(\cdot)$ of the associating random variable whose PDF holds zero-mean and unit variance. If additive noise assumed is spatially uncorrelated noise and absolutely uncorrelated from channels, it results in

$$\boldsymbol{\Sigma}_{xx}[n_T] = p[n_T] \boldsymbol{D}_a(\omega) \boldsymbol{B}(\sigma_\omega) \boldsymbol{D}_a^{\mathsf{H}}(\omega) + \sigma_{\mathrm{n}}^2 \boldsymbol{I} \quad (9)$$

where $p[n_T] \triangleq \rho |s[n_T]|^2$ stands for the total power observed at the sensor array. In what follows, we shall consider only the deterministic signal with constant modulus so that $\boldsymbol{\Sigma}_{xx}[n_T] = \boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta}_\circ)$; $\forall n_T$, where $\boldsymbol{\theta}_\circ$ is the true value of model parameter. Now suppose that based on the second-order statistic $\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta}_\circ)$ our problem is to find the nominal direction of arrival, ϕ, given the collected data $\boldsymbol{x}[n_T]; \forall n_T$, where true-valued parameter vector $\boldsymbol{\theta}_\circ \in \mathbb{R}^{4 \times 1}$ in the considered model can be defined by

$$\boldsymbol{\theta}_\phi \triangleq \begin{bmatrix} \phi & \sigma_\phi & p & \sigma_{\mathrm{n}}^2 \end{bmatrix}^{\mathsf{T}} \quad (10a)$$

$$\boldsymbol{\theta}_\omega \triangleq \begin{bmatrix} \omega & \sigma_\omega & p & \sigma_{\mathrm{n}}^2 \end{bmatrix}^{\mathsf{T}} \quad (10b)$$

for the physical and spatial frequency models, respectively. Let us introduce the matrix trace, derivative with respect to scalar and Krönecker product operator as $\lceil \mathbf{A} \rfloor$, $\dot{\mathbf{A}}(\mathrm{x}) \triangleq \frac{\partial}{\partial \mathrm{x}} \mathbf{A}(\mathrm{x})$ and $\otimes$. Under the central limit theorem, the snapshot data is also of Gaussianity with $\boldsymbol{x}[n_T] \sim \mathcal{N}_c(\boldsymbol{0}; \boldsymbol{\Sigma}_{xx}, \boldsymbol{O})$. To estimate the exact $\boldsymbol{\Sigma}_{xx}$, the sample covariance matrix $\hat{\boldsymbol{\Sigma}}_{xx} \in \mathbb{C}_{\mathbb{H}}^{N_E \times N_E}$ is given by

$$\hat{\boldsymbol{\Sigma}}_{xx} = \frac{1}{N_T} \sum_{n_T=1}^{N_T} \boldsymbol{x}[n_T] \boldsymbol{x}^{\mathsf{H}}[n_T]. \quad (11)$$

3 Separable Parameterizations

In this section, the column-stacking vectorization operator $\boldsymbol{v}_{\mathrm{c}}(\cdot)$ is performed to represent $\boldsymbol{\xi}_x \triangleq \mathcal{E}\langle \boldsymbol{x}^*[n_T] \otimes \boldsymbol{x}[n_T] \rangle = \boldsymbol{v}_{\mathrm{c}}(\boldsymbol{\Sigma}_{xx}) \in \mathbb{C}^{N_E^2 \times 1}$ in a certain parameterization.

3.1 Nominal Direction Parameterization

Let us define $\tilde{\boldsymbol{B}}(p, \sigma_\omega, \sigma_{\text{n}}^2) \triangleq p\boldsymbol{B}(\sigma_\omega) + \sigma_{\text{n}}^2 \boldsymbol{I} \in \mathbb{R}_{\mathbb{S},\mathbb{T}}^{N_E \times N_E}$. This exhibits a separable parameter $\boldsymbol{\vartheta}_\omega \in \mathbb{R}^{(N_E+1)\times 1}$ as

$$\boldsymbol{\vartheta} \triangleq \begin{bmatrix} \omega & \boldsymbol{\eta}_\omega^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} \tag{12}$$

where $\boldsymbol{\eta}_\omega \in \mathbb{R}^{N_E \times 1}$ is the first column vector in $\tilde{\boldsymbol{B}}(p, \sigma_\omega, \sigma_{\text{n}}^2)$. Such a parameterization results in

$$\boldsymbol{\xi}_x(\boldsymbol{\vartheta}_\omega) = \boldsymbol{\Omega}_\omega(\omega)\boldsymbol{\eta}_\omega \tag{13}$$

where full-rank matrix $\boldsymbol{\Omega}_\omega(\omega) : [-kd_E, kd_E] \mapsto \mathbb{C}_{\mathbb{F}}^{N_E^2 \times N_E}$ is $\boldsymbol{\Omega}_\omega(\omega) \triangleq \boldsymbol{\Phi}_a(\omega)\boldsymbol{\Xi}$ with full-rank binary selection matrix $\boldsymbol{\Xi} \in \mathbb{B}_{\mathbb{F}}^{N_E^2 \times N_E}$ corresponding to the Toeplitz structure of $\tilde{\boldsymbol{B}}(p, \sigma_\omega, \sigma_{\text{n}}^2)$ and nominal frequency parameterization matrix $\boldsymbol{\Phi}_a(\omega) \triangleq \boldsymbol{D}_a^{\mathsf{H}}(\omega) \otimes \boldsymbol{D}_a(\omega) : [-kd_E, kd_E] \mapsto \mathbb{C}_{\mathbb{D},\mathbb{U}}^{N_E^2 \times N_E^2}$. It was mentioned in [5] that based on the extended invariance principle the reparameterization between $\boldsymbol{\theta}_\omega$ and $\boldsymbol{\vartheta}_\omega$ yields the same performance.

3.2 Joint Parameterization of Nominal Direction and Angular Spread

Assume that we wish to joint estimate both ω and σ_ω. We must define $\boldsymbol{\eta}_{\omega,\sigma_\omega} \in \mathbb{R}^{2\times 1}$ as $\boldsymbol{\eta}_{\omega,\sigma_\omega} \triangleq \begin{bmatrix} p & \sigma_{\text{n}}^2 \end{bmatrix}$ so that

$$\boldsymbol{\xi}_x(\boldsymbol{\theta}_\omega) = \boldsymbol{\Omega}_{\omega,\sigma_\omega}(\omega, \sigma_\omega)\boldsymbol{\eta}_{\omega,\sigma_\omega} \tag{14}$$

where $\boldsymbol{\Omega}_{\omega,\sigma_\omega}(\omega, \sigma_\omega) : [-kd_E, kd_E] \times \mathbb{R}_+^{1\times 1} \mapsto \mathbb{C}_{\mathbb{F}}^{N_E^2 \times 2}$ is $\boldsymbol{\Omega}_{\omega,\sigma_\omega}(\omega, \sigma_\omega) \triangleq \begin{bmatrix} \boldsymbol{v}_{\text{c}}\left(\boldsymbol{D}_a(\omega)\boldsymbol{B}(\sigma_\omega)\boldsymbol{D}_a^{\mathsf{H}}(\omega)\right) & \boldsymbol{v}_{\text{c}}(\boldsymbol{I}) \end{bmatrix}$.

4 Decoupled AML Estimator

If we designate the nonparametric estimate $\hat{\boldsymbol{\Psi}}_{xx} \in \mathbb{C}_{\mathbb{H}}^{N_E^2 \times N_E^2}$ as $\hat{\boldsymbol{\Psi}}_{xx} \triangleq \hat{\boldsymbol{\Sigma}}_{xx}^{\mathsf{T}} \otimes \hat{\boldsymbol{\Sigma}}_{xx}$, then the AML nuisance estimate becomes [4]

$$\hat{\boldsymbol{\eta}}_{\text{AML}}(\boldsymbol{\iota}) = \left(\boldsymbol{\Omega}^{\mathsf{H}}(\boldsymbol{\iota})\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega}(\boldsymbol{\iota})\right)^{-1} \boldsymbol{\Omega}^{\mathsf{H}}(\boldsymbol{\iota})\hat{\boldsymbol{\Psi}}_{xx}^{-1}\hat{\boldsymbol{\xi}}_x. \tag{15}$$

Plugging the incomplete $\hat{\boldsymbol{\eta}}_{\text{AML}}(\boldsymbol{\iota})$ into $\boldsymbol{\xi}_x(\boldsymbol{\vartheta}_\omega) = \boldsymbol{\Omega}(\boldsymbol{\iota})\boldsymbol{\eta}$, we obtain

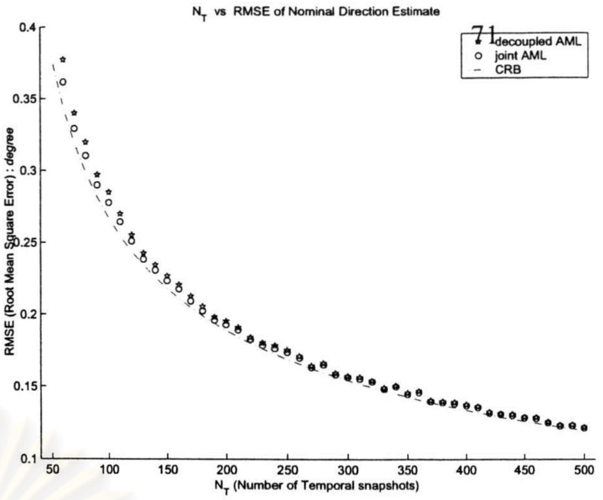
$$\boldsymbol{v}_{\text{c}}(\widehat{\boldsymbol{\Sigma}}_{xx}(\boldsymbol{\iota})) = \boldsymbol{\Omega}(\boldsymbol{\iota})\hat{\boldsymbol{\eta}}_{\text{AML}}(\boldsymbol{\iota}) \tag{16}$$

where $\widehat{\boldsymbol{\Sigma}}_{xx}(\boldsymbol{\iota}) \triangleq \boldsymbol{\Sigma}_{xx}(\boldsymbol{\iota}, \hat{\boldsymbol{\eta}}_{\text{AML}}(\boldsymbol{\iota}))$ is the concentrated covariance for AML estimate. Then, the AML estimator of the parameter of interest can be written as

$$\hat{\boldsymbol{\iota}}_{\text{AML}} = \arg\min_{\boldsymbol{\iota}} \ell_{\text{AML}}^{[N_T]}(\boldsymbol{\iota}) \tag{17a}$$

$$\ell_{\text{AML}}^{[N_T]}(\boldsymbol{\iota}) = \lceil \widehat{\boldsymbol{\Sigma}}_{xx}^{-1}(\boldsymbol{\iota})\hat{\boldsymbol{\Sigma}}_{xx} \rfloor + \ln|\widehat{\boldsymbol{\Sigma}}_{xx}(\boldsymbol{\iota})|. \tag{17b}$$

Now the question is implicitly imposed in what the parameter of interest, $\boldsymbol{\iota}$, should be. The following procedure enables us to an obvious answer for decoupled estimation



of nominal direction ϕ and its underlying angular spread σ_ϕ:

$$\boldsymbol{\iota} = \omega \overset{(13)}{\Rightarrow} \hat{\omega}_{\text{AML}} = \arg\min_{\omega} \ell_{\text{AML}}^{[N_T]}(\omega) \tag{18}$$

$$\boldsymbol{\iota} = \begin{bmatrix} \omega \\ \sigma_\omega \end{bmatrix} \overset{(14)}{\Rightarrow} \{\hat{\sigma}_\omega\}_{\text{AML}} = \arg\min_{\sigma_\omega} \ell_{\text{AML}}^{[N_T]}(\hat{\omega}_{\text{AML}}, \sigma_\omega). \tag{19}$$

Searching the minimum solution for $\hat{\omega}_{\text{AML}}$ and $\{\hat{\sigma}_\omega\}_{\text{AML}}$ according to two successive one-dimensional searches, we immediately obtain physical angle estimates via (6).

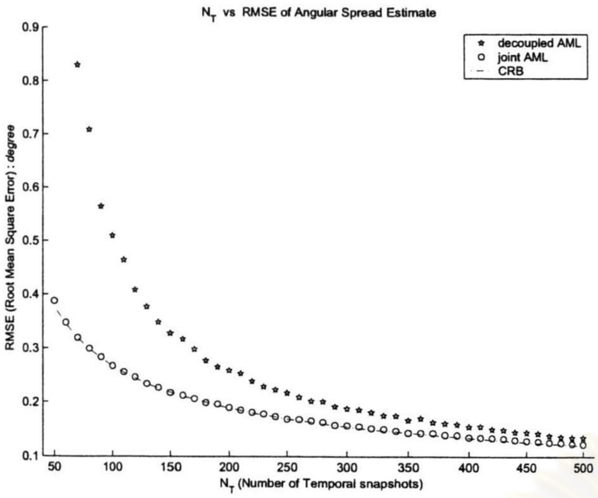
5 Numerical Examples

To demonstrate the impact of the proposed estimator, we commonly employ the ULA with half-wavelength separation to receive a QPSK (*q*uaternary *p*hase *s*hift *k*eying) signal whose strength are controllable with respect to noise variance by $\text{SNR} \triangleq 10\log\left(\frac{\sigma_{\text{s}}^2}{\sigma_{\text{n}}^2}\right)$. All significant parameters are set up, unless otherwise a variation on the parameter of interest will be specified individually in each figure, as the following table:

ϕ_o	σ_{ϕ_o}	σ_γ^2	SNR	N_P	N_E	N_R
$0°, 10°$	$5°$	0.01	10	100	8	1,000

Practically, the pseudo random number satisfied the Laplacian PDF $f_{\text{L}}(\delta_\phi|0,1)$ can be modified from $\delta_{\phi_{\text{L}}} = \frac{1}{\sqrt{2}}\ln\left(\frac{\delta_{\phi_{\text{U}}}}{\acute{\delta}_{\phi_{\text{U}}}}\right)$ [6] with any two independent uniform distributions $\delta_{\phi_{\text{U}}} \sim \mathcal{U}[0,1]$ and $\acute{\delta}_{\phi_{\text{U}}} \sim \mathcal{U}[0,1]$. Our empirical standard deviation is to average RMSE from a large number of independent runs (N_R).

Recently, it is shown that the AML estimator outperforms the WLS in non-asymptotic region [4]. Therefore, we shall investigate only the effect of decoupled estimation based on AML approach.



In Fig. 5, the joint AML estimator slightly outperforms the decoupled AML for small number of temporal snapshot. As shown in asymptotic performance assessment, both estimators achieve the CRB as the number of temporal snapshots tends to infinity.

For estimating the angular spread in Fig. 5, the decoupled AML estimator more deviates from the CRB than that shown in Fig. 5. This is because the second step for estimating the angular spread has imposed the uncertainty in nominal direction estimation. However, this effect will be gradually vanished when the nominal direction estimate is more accurate. In Fig. 5, it can be observed that both joint and decoupled AML estimations yields the same RMSE performance from large number of temporal snapshots.

6 Conclusion

A decoupled approach with two steps has been proposed for estimating the nominal direction and its underlying angular spread. It is intended to provide more numerical flexibility than the joint estimation in a certain application, *e.g.*, the situation where the angular spread might be not of interest in a while. Numerical simulation was also conducted to validate the asymptotic efficiency with respect to the joint estimation and the CRB. The numerical results are verified that the decoupled estimation can attain the CRB as same as in the joint estimation when employing a large number of temporal sanpshots.

References

[1] H. Krim and M. Viberg, Two decades of array signal processing research: The parametric approach, *IEEE Signal Processing Magazine,* 13, 1996, 67–94.

[2] T. Trump and B. Ottersten, Estimation of nominal direction of arrival and angular spread using an array of sensors, Technical Report, IR-S3-SB-9607, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, Apr. 1996. also appeared in *Signal Processing*, 50, 1996, 57–69.

[3] M. Bengtsson, Antenna array signal processing for high rank data models, Ph.D. Thesis, TRITA-S3-SB-9938, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, 1999.

[4] B. Täu Sieskul and S. Jitapunkul, An asymptotic maximum likelihood for joint estimation of nominal angles and angular spreads of multiple spatially distributed source, submitted, 2004.

[5] O. Besson and P. Stoica, Decoupled estimation of DOA and angular spread for a spatially distributed source, *IEEE Transactions on Signal Processing*, 48, 2000, 1872–1882.

[6] B. Täu Sieskul and S. Jitapunkul, Towards Laplacian angle deviation model for spatially distributed source localization, accepted to participate in *International Symposium on Communications and Information Technologies*, 2004.

Towards Laplacian Angle Deviation Model for Spatially Distributed Source Localization

Bamrung Täu Sieskul and Somchai Jitapunkul
Digital Signal Processing Research Laboratory, Department of Electrical Engineering, Chulalongkorn University, Bangkok Thailand, 10330, Tel: +662-0-2218-6900, Fax: +662-0-2218-6912, E-mail: bts@chula.com, somchai.j@chula.ac.th

Abstract— **This paper aims at presenting a more realistic Rayleigh channel model for the problem of finding the arrival direction in the presence of spatially spread source. Unlike the previous angular deviations assumed to be randomly distributed as uniform or normal (Gaussian), Laplacian assumption is herein taken into account instead of such so as to support recent field measurements. Regarding to the inherent accuracy imposed in our realistic model, we provide a rather systematic formulation of Cramér-Rao bound at hand and a derivative expression of Laplacian characteristic function aligned in the angular spread parameterization matrix. Incorporating maximum likelihood estimator into the model, some numerical results demonstrate that there exist small and considerable degradations of nominal angle and angular spread estimates owing to the heavier tail of Laplacian PDF.**

I. Introduction

Most works involved angular deviation model was assumed to be characterized as uniform or normal distribution in non-geometrical model [1]. However, it is not rather corresponding to outdoor field measurements, for instance, setting up in [2] and indoor experiments *e.g.*, in testbed [3]. For a certain application, the way to describe the angle characterization best is to find the most appropriate distribution for fitting the angle measurement result and a preassigned distribution with the smallest residual. In the model having to account for angular spread, this task has been carried out by means of *goodness of fit* in [4] and [5]. It was argued that the Laplacian distribution yields the best match in both urban and rural areas even in a non-LOS situation. Consequently, a lot of system limitation analysis recently focuses on Laplacian power density (see *e.g.*, [6], [7] and references therein).

Interestingly enough, companion studies dealing with either geometrical or non-geometrical models were provided without concerning the Laplacian power density in [8]. Moreover, it was conducted by invoking subspace-based estimators which is biased and eventually sub-optimal in high rank data model [9].

Here we proceed on spatially distributed source localization model by replacing the existing angle PDFs in many literatures with the Laplacian distribution. In the aspect of parameter estimation, our problem belongs to jointly estimating the nominal angle and its underlying angular spread from which random perturbation is of Laplacian random numbers generated by a simple method. Not only pure simulation is insightful, but achievable performance of the Laplacian model with respect to two other types of angular distribution is also investigated theoretically via a variational formulation of the *C*ramér-*R*ao *b*ound (CRB).

The rest of this paper is arranged as follows: section II reviews the model of passive sensor array that employs the *u*niform *l*inear array (ULA) to receive a lot of distributed source signals. In section III, an alternative CRB formulation is derived. Certainly, in section IV the Laplacian distribution for source localization is stated without realizing truncated range. To validate the applicability of proposed deviation model, numerical examples demonstrate some reflections and robustness of more realistic assumption in section V. Finally, the summarization at section VI is given for all content in the paper.

II. Spatially Distributed Source Model

Restrict our attention to a number of signals transmitting through a channel and then impinging on the sensor array antenna. With phase reference at the first element (*f*irst *e*lement *r*eference (FER)), the array response vector $\boldsymbol{a}(\phi) : [-90°, 90°] \mapsto \mathbb{C}^{N_E \times 1}$ can be, in general, written ideally as

$$\boldsymbol{a}(\phi) \triangleq \begin{bmatrix} 1 & e^{\imath k d_E \sin(\phi)} & \cdots & e^{\imath k d_E (N_E - 1)\sin(\phi)} \end{bmatrix}^{\mathsf{T}} \quad (1)$$

where $k = \frac{2\pi}{\lambda}$ designates the wave number and N_E is the number of sensor elements. In according with linear regression analysis, the array output at time instant n_T can be characterized by the snapshot $\boldsymbol{x}[n_T] \in \mathbb{C}^{N_E \times 1}$ in a flat fading channel such that [13, p. 25]

$$\boldsymbol{x}[n_T] = \boldsymbol{H}[n_T]\boldsymbol{s}[n_T] + \boldsymbol{n}[n_T] \quad (2)$$

where N_S denotes the number of source signals. Here the time-varying channel matrix $\boldsymbol{H}[n_T] \in \mathbb{C}^{N_E \times N_S}$ and the source signal vector $\boldsymbol{s}[n_T] \in \mathbb{C}^{N_S \times 1}$ are collected as

$$\boldsymbol{H}[n_T] \triangleq \begin{bmatrix} \boldsymbol{h}_1[n_T] & \boldsymbol{h}_2[n_T] & \cdots & \boldsymbol{h}_{N_S}[n_T] \end{bmatrix} \quad (3a)$$

$$\boldsymbol{s}[n_T] \triangleq \begin{bmatrix} s_1[n_T] & s_2[n_T] & \cdots & s_{N_S}[n_T] \end{bmatrix}^{\mathsf{T}} \quad (3b)$$

where $\boldsymbol{h}_{n_S}[n_T] \in \mathbb{C}^{N_E \times 1}$ and $\boldsymbol{n}[n_T] \in \mathbb{C}^{N_E \times 1}$ signify the channel vector and additive noise. As previously developed, most local scattering models assume that the nominal angle ϕ is deterministic while the angular deviation δ_ϕ and the associating path gain γ are considered as stochastic quantities eventually. In the presence of local scattering, all $N_P[n_S]$ non-*l*ine-*o*f-*s*ight (NLOS) paths with i.i.d. (*i*dentical and *i*ndependent *d*istribution) are, under the superposition theorem (see *e.g.* [10], [11], [12] and [13])

$$\boldsymbol{h}_{n_S}[n_T] = \sum_{n_P=1}^{N_P[n_S]} \gamma_{n_P,n_S}[n_T]\boldsymbol{a}(\phi_{n_S} + \delta_{\phi_{n_P,n_S}}[n_T]). \quad (4)$$

For taking an incoherently distributed channel [10] into account, the second order statistic of a certain incoming ray yields [1]

$$\mathcal{E}\left\langle \gamma_{n_P,n_S}[n_T]\gamma^*_{\acute{n}_P,\acute{n}_S}[\acute{n}_T]\right\rangle = \sigma^2_{\gamma_{n_S}}\delta_{n_P,\acute{n}_P}\delta_{n_S,\acute{n}_S}\delta_{n_T,\acute{n}_T} \quad (5)$$

where $\delta_{\bullet,\bullet}$ signifies the Krönecker delta function, σ^2_γ designates the power due to any path. Accounting for a large number of incoming rays, the channel vector seems, under the central limit theorem, to be circularly-symmetric Gaussian distribution, *i.e.*, $\boldsymbol{h}_{n_S}[n_T] \sim \mathcal{N}_c\left(\boldsymbol{0}; \boldsymbol{\Sigma}_{hh}, \boldsymbol{O}\right)$ with the channel covariance $\boldsymbol{\Sigma}_{hh} \triangleq \mathcal{E}\left\langle \boldsymbol{h}[n_T]\boldsymbol{h}^{\mathsf{H}}[n_T]\right\rangle \in \mathbb{C}_{\mathbb{H}}^{N_E\times N_E}$ given by

$$\boldsymbol{\Sigma}_{hh}(\rho,\phi,\sigma_\phi) \approx \rho \int f(\delta_\phi|0;\sigma^2_\phi)\boldsymbol{a}(\phi+\delta_\phi)\boldsymbol{a}^{\mathsf{H}}(\phi+\delta_\phi)d\delta_\phi \quad (6)$$

where $\rho_{n_S} \triangleq N_P[n_S]\sigma^2_\gamma$ signifies the cluster power due to all paths and $f(\delta_\phi|0;\sigma^2_\phi)$ denotes the conditional PDF for random deviation δ_ϕ given a priori knowledge of the angular spread σ_ϕ. As encountered, a family of symmetric distributions with zero mean and variance σ^2_ϕ is in most modelled as

$$f(\delta_\phi|0;\sigma^2_\phi) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_\phi} e^{-\frac{1}{2}\frac{\delta^2_\phi}{\sigma^2_\phi}} & ; \text{Gaussian} \\ \frac{1}{2\sqrt{3}\sigma_\phi} \sqcap\left[-\sqrt{3}\sigma_\phi, \sqrt{3}\sigma_\phi\right] & ; \text{Uniform.} \end{cases} \quad (7)$$

In instead of such physical angles ϕ and σ_ϕ, the spatial frequency response is preferable due to the better accuracy of approximating the first-order Taylor series around the array broadside [13]. Introduce a parameterization of the spatial frequency ω and its associating standard deviation σ_ω as

$$\omega(\phi) = kd_E \sin(\phi) \quad (8a)$$
$$\sigma_\omega(\phi,\sigma_\phi) = kd_E \cos(\phi)\sigma_\phi. \quad (8b)$$

For small angular spreads, the so-called *spatial frequency* approximation results in a separable form of

$$\boldsymbol{\Sigma}_{hh}(\rho,\omega,\sigma_\omega) \simeq \rho\boldsymbol{D}_a(\omega)\boldsymbol{B}(\sigma_\omega)\boldsymbol{D}_a^{\mathsf{H}}(\omega) \quad (9)$$

where the $(n_E, \acute{n}_E)$-th elements of diagonal and unitary matrix $\boldsymbol{D}_a(\omega) : [-kd_E, kd_E] \mapsto \mathbb{C}_{\mathbb{D},\mathbb{U}}^{N_E\times N_E}$ and symmetric Toeplitz matrix $\boldsymbol{B}(\sigma_\omega) : [0, kd_E\sigma_\phi] \mapsto \mathbb{R}_{\mathbb{S},\mathbb{T}}^{N_E\times N_E}$ are given from [13, p. 22]

$$[\boldsymbol{D}_a(\omega)]_{[n_E,\acute{n}_E]} = e^{\imath(n_E-1)\omega}\delta_{n_E,\acute{n}_E} \quad (10a)$$
$$[\boldsymbol{B}(\sigma_\omega)]_{[n_E,\acute{n}_E]} = f_{\mathcal{F}}((n_E-\acute{n}_E)\sigma_\omega|0;1) \quad (10b)$$

with the characteristic function $f_{\mathcal{F}}(t|,0;1) \triangleq \mathcal{F}(f(\delta_\omega|,0;1))$ of associating random variable whose PDF holds zero-mean and unit variance. In a certain situation, the $(n_E, \acute{n}_E)$-th entry in $\boldsymbol{B}(\sigma_\omega)$ can be expressed as [13, p. 28]

$$[\boldsymbol{B}(\sigma_\omega)]_{[n_E,\acute{n}_E]} = \begin{cases} e^{-\frac{1}{2}(n_E-\acute{n}_E)^2\sigma^2_\omega} & ; \text{Gaussian} \\ \frac{\sin((n_E-\acute{n}_E)\sqrt{3}\sigma_\omega)}{(n_E-\acute{n}_E)\sqrt{3}\sigma_\omega} & ; \text{Uniform.} \end{cases} \quad (11)$$

If additive noise assumed is spatially uncorrelated noise and absolutely uncorrelated from channels, it indeed yields

$$\boldsymbol{\Sigma}_{xx}[n_T] = \sum_{n_S=1}^{N_S} p_{n_S}[n_T]\boldsymbol{D}_a(\omega_{n_S})\boldsymbol{B}(\sigma_{\omega_{n_S}})\boldsymbol{D}_a^{\mathsf{H}}(\omega_{n_S}) + \sigma^2_n\boldsymbol{I} \quad (12)$$

where $p_{n_S}[n_T] \triangleq \rho_{n_S}|s_{n_S}[n_T]|^2$ stands for the total power observed at the sensor array due to the n_Sth signal. In what follows, we shall consider only the deterministic signal with constant modulus so that $\boldsymbol{\Sigma}_{xx}[n_T] = \boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta}_\circ)\ ; \forall n_T$, where $\boldsymbol{\theta}_\circ$ signifies the true value of model parameter. Let the sample covariance matrix $\hat{\boldsymbol{\Sigma}}_{xx} \in \mathbb{C}_{\mathbb{H}}^{N_E\times N_E}$ be

$$\hat{\boldsymbol{\Sigma}}_{xx} = \frac{1}{N_T}\sum_{n_T=1}^{N_T}\boldsymbol{x}[n_T]\boldsymbol{x}^{\mathsf{H}}[n_T]. \quad (13)$$

Now suppose that our problem is to find the directions of arrival $\boldsymbol{\phi}$ given the collected data $\boldsymbol{x}[n_T]; \forall n_T$ based on the second-order statistic $\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta}_\circ)$. The model parameter vector $\boldsymbol{\theta}_\circ \in \mathbb{R}^{(3N_S+1)\times 1}$ is in nature defined by

$$\boldsymbol{\theta}_\phi \triangleq \left[\boldsymbol{\phi}^{\mathsf{T}} \quad \boldsymbol{\sigma}_\phi^{\mathsf{T}} \quad \boldsymbol{p}^{\mathsf{T}} \quad \sigma^2_n\right]^{\mathsf{T}} \quad (14a)$$
$$\boldsymbol{\theta}_\omega \triangleq \left[\boldsymbol{\omega}^{\mathsf{T}} \quad \boldsymbol{\sigma}_\omega^{\mathsf{T}} \quad \boldsymbol{p}^{\mathsf{T}} \quad \sigma^2_n\right]^{\mathsf{T}} \quad (14b)$$

for respectively the physical and spatial frequency models.

III. A Formulation of Cramér-Rao Bound

Let us introduce the matrix trace and derivative with respect to scalar as $\lceil\mathbf{A}\rfloor$ and $\dot{\mathbf{A}}(\mathrm{x}) \triangleq \frac{\partial}{\partial \mathrm{x}}\mathbf{A}(\mathrm{x})$, respectively. Under the central limit theorem, the n_Tth snapshot is also of circularly-symmetric complex-valued Gaussianity with $\boldsymbol{x}[n_T] \sim \mathcal{N}_c\left(\boldsymbol{0}; \boldsymbol{\Sigma}_{xx}, \boldsymbol{O}\right)$. It was proposed in [11] that the *m*aximum *l*ikelihood (ML) estimator of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}}_{\mathrm{ML}} = \arg\min_{\boldsymbol{\theta}}\left(\lceil\boldsymbol{\Sigma}^{-1}_{xx}(\boldsymbol{\theta})\hat{\boldsymbol{\Sigma}}_{xx}\rfloor + \ln|\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta})|\right) \quad (15)$$

Recall the Slepian-Bangs's formula accounting for the zero-mean random vector $\boldsymbol{x}[n_T]$. This leads to the $(n, \acute{n})$th element of *F*isher *i*nformation *m*atrix (FIM) [14]

$$[\boldsymbol{I}_{\mathrm{F}}(\boldsymbol{\theta})]_{[n,\acute{n}]} = N_T\lceil\boldsymbol{\Sigma}^{-1}_{xx}\dot{\boldsymbol{\Sigma}}_{xx}(\theta_n)\boldsymbol{\Sigma}^{-1}_{xx}\dot{\boldsymbol{\Sigma}}_{xx}(\theta_{\acute{n}})\rfloor \quad (16)$$

where the scalars θ_n and $\theta_{\acute{n}}$ are the nth and $\acute{n}$th elements of the parameter vector $\boldsymbol{\theta}$ for indices $n, \acute{n} \in \{1, 2, \ldots, 3N_S+1\}$. Let the derivative matrix $\nabla_\theta(\boldsymbol{\xi}_x) \in \mathbb{C}^{N_E^2\times(3N_S+1)}$ of the vector $\boldsymbol{\xi}_x(\boldsymbol{\theta}) \triangleq \boldsymbol{v}_{\mathrm{c}}\left(\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta})\right) : \mathbb{R}^{(3N_S+1)\times 1} \mapsto \mathbb{C}^{N_E^2\times 1}$ be given by

$$\nabla_\theta(\boldsymbol{\xi}_x) \triangleq \frac{\partial}{\partial\boldsymbol{\theta}^{\mathsf{T}}}\boldsymbol{\xi}_x(\boldsymbol{\theta}) = \left[\dot{\boldsymbol{\xi}}_x(\theta_1) \quad \cdots \quad \dot{\boldsymbol{\xi}}_x(\theta_{3N_S+1})\right] \quad (17)$$

where $\boldsymbol{v}_{\mathrm{c}}(\cdot)$ denotes column-stacking vectorization operator. One can fulfill the FIM (see *e.g.* [15], [16], [17] and [18]) as

$$\boldsymbol{I}_{\mathrm{F}}(\boldsymbol{\theta}) = N_T \nabla_{\theta}^{\mathsf{H}}(\boldsymbol{\xi}_x)\boldsymbol{\Psi}_{xx}^{-1}(\boldsymbol{\theta})\nabla_{\theta}(\boldsymbol{\xi}_x) \tag{18}$$

where $\boldsymbol{\Psi}_{xx}(\boldsymbol{\theta}) \triangleq \boldsymbol{\Sigma}_{xx}^{\mathsf{T}}(\boldsymbol{\theta})\otimes\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta})\in\mathbb{C}_{\mathbb{H}}^{N_E^2\times N_E^2}$. Based on the spatial frequency approximation, $\boldsymbol{\xi}_x(\boldsymbol{\theta}_\omega)$ can be justified as

$$\boldsymbol{\xi}_x(\boldsymbol{\theta}_\omega) = \boldsymbol{\Omega}(\boldsymbol{\omega},\boldsymbol{\sigma}_\omega)\boldsymbol{\eta}(\boldsymbol{p},\sigma_n^2) \tag{19}$$

whence $\boldsymbol{\Omega}(\boldsymbol{\omega},\boldsymbol{\sigma}_\omega):\mathbb{R}^{(2N_S)\times 1}\mapsto\mathbb{C}_{\mathbb{F}}^{N_E^2\times(N_S+1)}$, defined by

$$\begin{aligned}&\boldsymbol{\Omega}(\boldsymbol{\omega},\boldsymbol{\sigma}_\omega)\\ &\triangleq\Big[\boldsymbol{\Phi}_a(\omega_1)\boldsymbol{\Xi}\boldsymbol{b}(\sigma_{\omega_1})\ \ \cdots\ \ \boldsymbol{\Phi}_a(\omega_{N_S})\boldsymbol{\Xi}\boldsymbol{b}(\sigma_{\omega_{N_S}})\ \ \boldsymbol{v}_{\mathrm{c}}(\boldsymbol{I})\Big]\end{aligned}$$

and $\boldsymbol{\eta}(\boldsymbol{p},\sigma_n^2)\triangleq\begin{bmatrix}\boldsymbol{p}^{\mathsf{T}} & \sigma_n^2\end{bmatrix}^{\mathsf{T}}:\mathbb{R}_+^{(N_S+1)\times 1}\mapsto\mathbb{R}_+^{(N_S+1)\times 1}$ constitute a new parameterization due to binary selection matrix of full rank $\boldsymbol{\Xi}\in\mathbb{B}^{N_E^2\times N_E}$ and nominal frequency matrix $\boldsymbol{\Phi}_a(\omega)\triangleq\boldsymbol{D}_a^{\mathsf{H}}(\omega)\otimes\boldsymbol{D}_a(\omega):[-kd_E,kd_E]\mapsto\mathbb{C}_{\mathbb{D},\mathbb{U}}^{N_E^2\times N_E^2}$. Straightforward calculating the derivatives, we obtain

$$\dot{\boldsymbol{\xi}}_x(\omega_{n_S}) = p_{n_S}\dot{\boldsymbol{\Phi}}_a(\omega_{n_S})\boldsymbol{\Xi}\boldsymbol{b}(\sigma_{\omega_{n_S}}) \tag{20a}$$

$$\dot{\boldsymbol{\xi}}_x(\sigma_{\omega_{n_S}}) = p_{n_S}\boldsymbol{\Phi}_a(\omega_{n_S})\boldsymbol{\Xi}\dot{\boldsymbol{b}}(\sigma_{\omega_{n_S}}) \tag{20b}$$

where two derivative results are adaptable from

$$\dot{\boldsymbol{\Phi}}_a(\omega) = \Big(\boldsymbol{D}_a^{\mathsf{H}}(\omega)\otimes\dot{\boldsymbol{D}}_a(\omega)\Big)+\Big(\dot{\boldsymbol{D}}_a^{\mathsf{H}}(\omega)\otimes\boldsymbol{D}_a(\omega)\Big) \tag{21a}$$

$$\dot{\boldsymbol{b}}(\sigma_\omega) = \Big[\dot{\boldsymbol{B}}(\sigma_\omega)\Big]_{[:,1]}. \tag{21b}$$

It is easy to see that $\dot{\boldsymbol{D}}_a(\omega)=\imath\boldsymbol{\Lambda}\boldsymbol{D}_a(\omega)$, where $[\boldsymbol{\Lambda}]_{[n,\acute{n}]}=n\delta_{n,\acute{n}}\,;n\in\{0,1,\ldots,N_E-1\}$ is a diagonal matrix. Relying on each PDF, we encounter

$$\Big[\dot{\boldsymbol{B}}(\sigma_\omega)\Big]_{[n_E,\acute{n}_E]} = \begin{cases}-(n_E-\acute{n}_E)^2\sigma_\omega[\boldsymbol{B}(\sigma_\omega)]_{[n_E,\acute{n}_E]} & ;\text{Gaussian}\\ \frac{1}{\sigma_\omega}\Big(\cos((n_E-\acute{n}_E)\sqrt{3}\sigma_\omega)-[\boldsymbol{B}(\sigma_\omega)]_{[n_E,\acute{n}_E]}\Big) & ;\text{Uniform.}\end{cases} \tag{22}$$

By accumulating all derivatives into $\nabla_\omega(\boldsymbol{\xi}_x)\triangleq\frac{\partial}{\partial\boldsymbol{\omega}^{\mathsf{T}}}\boldsymbol{\xi}_x(\boldsymbol{\omega})\in\mathbb{C}^{N_E^2\times N_S}$ and $\nabla_{\sigma_\omega}(\boldsymbol{\xi}_x)\triangleq\frac{\partial}{\partial\boldsymbol{\sigma}_\omega^{\mathsf{T}}}\boldsymbol{\xi}_x(\boldsymbol{\sigma}_\omega)\in\mathbb{C}^{N_E^2\times N_S}$ as

$$\nabla_\omega(\boldsymbol{\xi}_x) = \Big[\dot{\boldsymbol{\xi}}_x(\omega_1)\ \ \dot{\boldsymbol{\xi}}_x(\omega_2)\ \ \cdots\ \ \dot{\boldsymbol{\xi}}_x(\omega_{N_S})\Big] \tag{23a}$$

$$\nabla_{\sigma_\omega}(\boldsymbol{\xi}_x) = \Big[\dot{\boldsymbol{\xi}}_x(\sigma_{\omega_1})\ \ \dot{\boldsymbol{\xi}}_x(\sigma_{\omega_2})\ \ \cdots\ \ \dot{\boldsymbol{\xi}}_x(\sigma_{\omega_{N_S}})\Big] \tag{23b}$$

the derivative $\nabla_{\vartheta_\omega}(\boldsymbol{\xi}_x)\triangleq\frac{\partial}{\partial\boldsymbol{\vartheta}_\omega^{\mathsf{T}}}\boldsymbol{\xi}_x(\boldsymbol{\vartheta})\in\mathbb{C}^{N_E^2\times(2N_S)}$ with respect to $\boldsymbol{\vartheta}_\omega\triangleq[\boldsymbol{\omega}^{\mathsf{T}}\ \ \boldsymbol{\sigma}_\omega^{\mathsf{T}}]^{\mathsf{T}}\in\mathbb{R}^{(2N_S)\times 1}$ can be represented by

$$\nabla_{\vartheta_\omega}(\boldsymbol{\xi}_x) = \big[\nabla_\omega(\boldsymbol{\xi}_x)\ \ \nabla_{\sigma_\omega}(\boldsymbol{\xi}_x)\big] \tag{24}$$

Then, the Jacobian $\nabla_\eta(\boldsymbol{\xi}_x)\triangleq\frac{\partial}{\partial\boldsymbol{\eta}^{\mathsf{T}}}\boldsymbol{\xi}_x(\boldsymbol{\eta})=\boldsymbol{\Omega}(\boldsymbol{\omega},\boldsymbol{\sigma}_\omega)$ results in

$$\nabla_{\theta_\omega}(\boldsymbol{\xi}_x) = \big[\nabla_{\vartheta_\omega}(\boldsymbol{\xi}_x)\ \ \boldsymbol{\Omega}(\boldsymbol{\omega},\boldsymbol{\sigma}_\omega)\big]. \tag{25}$$

Since one can write $\boldsymbol{\theta}_\omega(\boldsymbol{\theta}_\phi)$, the Jacobian matrix $\boldsymbol{J}_{\theta_\phi}(\boldsymbol{\theta}_\omega)\triangleq\frac{\partial}{\partial\boldsymbol{\theta}_\phi^{\mathsf{T}}}\boldsymbol{\theta}_\omega(\boldsymbol{\theta}_\phi)\in\mathbb{R}^{(3N_S+1)\times(3N_S+1)}$ allows the CRB derived from spatial frequency model to be transformed into the physical model via[1] [19, pp. 45–46]

$$\boldsymbol{J}_{\theta_\phi}(\boldsymbol{\theta}_\omega) = \begin{bmatrix}\boldsymbol{J}_{\vartheta_\phi}(\boldsymbol{\vartheta}_\omega) & \boldsymbol{O} & \boldsymbol{0}\\ \boldsymbol{O}^{\mathsf{T}} & \boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0}^{\mathsf{T}} & \boldsymbol{0}^{\mathsf{T}} & 1\end{bmatrix}. \tag{26}$$

Furthermore, each element of the significant block matrix $\boldsymbol{J}_{\vartheta_\phi}(\boldsymbol{\vartheta}_\omega)\triangleq\frac{\partial}{\partial\boldsymbol{\vartheta}_\phi^{\mathsf{T}}}\boldsymbol{\vartheta}_\omega(\boldsymbol{\theta}_\phi)\in\mathbb{R}^{(2N_S)\times(2N_S)}$ can be drawn from

$$\begin{aligned}[\boldsymbol{J}_{\vartheta_\phi}(\boldsymbol{\vartheta}_\omega)]_{[n_S,n_S]} &= \dot{\omega}_{n_S}(\phi_{n_S})\\ &= kd_E\cos(\phi_{n_S})\end{aligned} \tag{27a}$$

$$\begin{aligned}[\boldsymbol{J}_{\vartheta_\phi}(\boldsymbol{\vartheta}_\omega)]_{[N_S+n_S,n_S]} &= \dot{\sigma}_{\omega_{n_S}}(\phi_{n_S})\\ &= -kd_E\sigma_{\phi_{n_S}}\sin(\phi_{n_S})\end{aligned} \tag{27b}$$

$$\begin{aligned}[\boldsymbol{J}_{\vartheta_\phi}(\boldsymbol{\vartheta}_\omega)]_{[N_S+n_S,N_S+n_S]} &= \dot{\sigma}_{\omega_{n_S}}(\sigma_{\phi_{n_S}})\\ &= kd_E\cos(\phi_{n_S}).\end{aligned} \tag{27c}$$

The Jacobian matrix $\nabla_{\theta_\phi}(\boldsymbol{\xi}_\circ)$ therefore becomes

$$\nabla_{\theta_\phi}(\boldsymbol{\xi}_\circ) = \nabla_{\theta_\omega}(\boldsymbol{\xi}_\circ)\boldsymbol{J}_{\theta_\phi}(\boldsymbol{\theta}_\omega). \tag{28}$$

In conjunction with $\boldsymbol{B}_{\mathrm{CR}(\hat{\boldsymbol{\theta}})}(\boldsymbol{\theta}_\circ)\triangleq\boldsymbol{I}_{\mathrm{F}}^{-1}(\boldsymbol{\theta}_\circ)$ [19], the CRB matrix accounted for estimating $\boldsymbol{\theta}_\phi$ is at last given by

$$\boldsymbol{B}_{\mathrm{CR}(\hat{\boldsymbol{\theta}}_\phi)}(\boldsymbol{\theta}_\circ) = \frac{1}{N_T}\Big(\nabla_{\theta_\phi}^{\mathsf{H}}(\boldsymbol{\xi}_\circ)\boldsymbol{\Psi}(\boldsymbol{\theta}_\circ)\nabla_{\theta_\phi}(\boldsymbol{\xi}_\circ)\Big)^{-1} \tag{29}$$

where $\nabla_{\theta_\phi}(\boldsymbol{\xi}_\circ)$ denotes $\frac{\partial}{\partial\boldsymbol{\theta}_\phi^{\mathsf{T}}}\boldsymbol{\xi}_x(\boldsymbol{\theta}_\phi)|_{\boldsymbol{\theta}_\phi=\boldsymbol{\theta}_\circ}$.

IV. Laplacian Angle Deviation Models

Next we shall provide a more realistic model for reflecting relevant channel features precisely.

A. Laplacian Distribution with Infinite Range

The Laplacian distribution with infinite range $\delta_\phi\in(-\infty,\infty)$ is customarily expressed by the PDF [20, p. 166]

$$f_{\mathrm{L}}(\delta_\phi|\bar{\delta}_\phi;\sigma_\phi^2) = \frac{1}{\sqrt{2}\sigma_\phi}e^{-\frac{1}{\sigma_\phi}\sqrt{2}|\delta_\phi-\bar{\delta}_\phi|}\quad;-\infty<\delta_\phi<\infty \tag{30}$$

with mean $\bar{\delta}_\phi\triangleq\mathcal{E}\langle\delta_\phi\rangle$ and variance $\sigma_\phi^2\triangleq\mathcal{E}\langle(\delta_\phi-\bar{\delta}_\phi)^2\rangle$. Taking the Fourier transform, it yields [21, p. 930]

$$\mathcal{F}(f_{\mathrm{L}}(\delta_\phi|\bar{\delta}_\phi;\sigma_\phi^2)) = \frac{2}{2+\sigma_\phi^2t^2}e^{\imath\bar{\delta}_\phi t}. \tag{31}$$

For the standard Laplacian distribution, this leads to

$$\mathcal{F}(f_{\mathrm{L}}(\delta_\phi|0;1)) = \frac{2}{2+t^2} \tag{32}$$

which equals the characteristic function in [22, p. 398].

[1]This is available from $J_{\vartheta_\phi}(\boldsymbol{p})\triangleq\frac{\partial}{\partial\boldsymbol{\vartheta}_\phi^{\mathsf{T}}}\boldsymbol{p}=\boldsymbol{O}^{\mathsf{T}}$, $J_p(\boldsymbol{\vartheta}_\phi)\triangleq\frac{\partial}{\partial\boldsymbol{p}^{\mathsf{T}}}\boldsymbol{\vartheta}_\phi=\boldsymbol{O}$, $J_p(\boldsymbol{\vartheta}_\phi)\triangleq\frac{\partial}{\partial\boldsymbol{p}^{\mathsf{T}}}\boldsymbol{p}=\boldsymbol{I}_{(N_S)}$, $\frac{\partial}{\partial\boldsymbol{\vartheta}_\phi^{\mathsf{T}}}\sigma_n^2=\boldsymbol{0}^{\mathsf{T}}$, $\frac{\partial}{\partial\sigma_n^2}\boldsymbol{\vartheta}_\phi=\boldsymbol{0}$, $\frac{\partial}{\partial\boldsymbol{p}^{\mathsf{T}}}\sigma_n^2=\boldsymbol{0}^{\mathsf{T}}$, $\frac{\partial}{\partial\sigma_n^2}\boldsymbol{p}=\boldsymbol{0}$ and $\frac{\partial}{\partial\sigma_n^2}\sigma_n^2=1$.

B. Laplacian Distribution with Truncated Range

If the random variable δ_ϕ is represented as spatial deviation angle, we always encounter restriction of $|\delta_\phi - \bar{\delta}_\phi| \leq \pi$ [7]. Regarding to both edges, we arrive at (see *e.g.* [4] and [5])

$$f_{\mathrm{TL}}(\delta_\phi|\bar{\delta}_\phi;\sigma_\phi^2) = \frac{c_{\mathrm{L}}(\sigma_\phi)}{\sqrt{2}\sigma_\phi}e^{-\frac{1}{\sigma_\phi}\sqrt{2}|\delta_\phi-\bar{\delta}_\phi|} \quad ;\bar{\delta}_\phi-\pi \leq \delta_\phi < \bar{\delta}_\phi+\pi \tag{33}$$

where $c_{\mathrm{L}}(\sigma_\phi)$, given by $c_{\mathrm{L}}(\sigma_\phi) = \frac{1}{1-e^{-\frac{1}{\sigma_\phi}\sqrt{2}\pi}}$ (see *e.g.* [6]), is a constant to normalize the truncated PDF a density function. In according with standard PDF, it then remains

$$\mathcal{F}(f_{\mathrm{TL}}(\delta_\phi|0;1)) = \frac{2}{(1-e^{-\sqrt{2}\pi})(2+t^2)}. \tag{34}$$

C. Laplacian Distribution for Source Localization

It is noteworthy that to be $\mathcal{F}(f_{\mathrm{TL}}(\delta_\phi|0;1))$, $\mathcal{F}(f_{\mathrm{L}}(\delta_\phi|0;1))$ must be scaled by such a scalar $c_{\mathrm{L}}(1)$. Henceforth, whether angular truncation is of interest or not, our source localization model will not make a matter because the scalar does not depend on model parameter $\boldsymbol{\theta}$.

Proposition 1: For computational simplicity, we make use of the Laplacian PDF with zero-mean and infinite range, given by

$$f(\delta_\phi|0;\sigma_\phi^2) = \frac{1}{\sqrt{2}\sigma_\phi}e^{-\frac{1}{\sigma_\phi}\sqrt{2}|\delta_\phi|}. \tag{35}$$

In term of spatial frequency model, it follows that

$$[\boldsymbol{B}(\sigma_\omega)]_{[n_E,\acute{n}_E]} = \frac{1}{1+\frac{1}{2}(n_E-\acute{n}_E)^2\sigma_\omega^2}. \tag{36}$$

As required in CRB calculation, the derivative of spatial fading correlation becomes

$$[\dot{\boldsymbol{B}}(\sigma_\omega)]_{[n_E,\acute{n}_E]} = -(n_E-\acute{n}_E)^2\sigma_\omega[\boldsymbol{B}(\sigma_\omega)]^2_{[n_E,\acute{n}_E]}. \tag{37}$$

V. Numerical Examples

To demonstrate the impact of the proposed deviation model, we commonly employ the ULA with half-wavelength separation and *f*irst-*e*lement *r*eference (FER) to receive a QPSK (*q*uaternary *p*hase *s*hift *k*eying) signal whose strength is controlled with respect to noise variance by

$$\mathrm{SNR} \triangleq 10\log\left(\frac{\sigma_s^2}{\sigma_n^2}\right). \tag{38}$$

All significant parameters are set up as the following table:

$\phi_\circ$	$\sigma_{\phi_\circ}$	σ_γ^2	SNR	N_S	N_P	N_E
0°	5°	0.001	10	1	1,000	8

As modified from [23, p. 94], the pseudo random number satisfied the Laplacian PDF $f_{\mathrm{L}}(\delta_\phi|0;1)$ can be generated by

$$\delta_{\phi_{\mathrm{L}}} = \frac{1}{\sqrt{2}}\ln\left(\frac{\delta_{\phi_{\mathrm{U}}}}{\acute{\delta}_{\phi_{\mathrm{U}}}}\right) \tag{39}$$

with any independent $\delta_{\phi_{\mathrm{U}}} \sim \mathcal{U}[0,1]$ and $\acute{\delta}_{\phi_{\mathrm{U}}} \sim \mathcal{U}[0,1]$. Our empirical standard deviation (*r*oot *m*ean *s*quare *e*rror (RMSE)) is to calculate the square root of averaging the squared error from a large number of realizations (N_R).

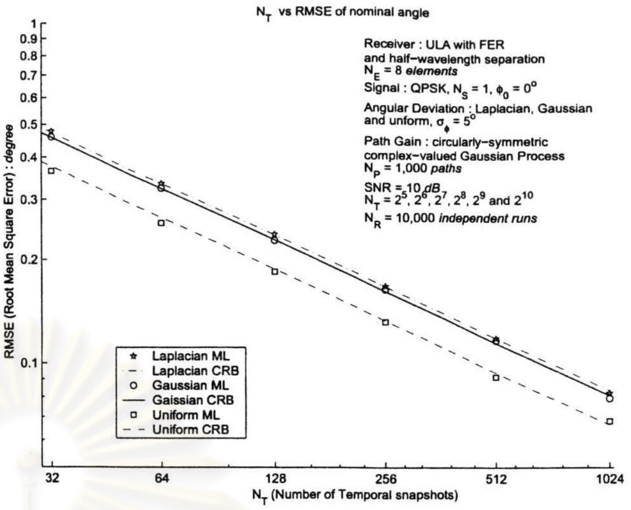


Fig. 1. RMSE of the nominal angle ϕ in three individual scenarios with appropriate assumptions.

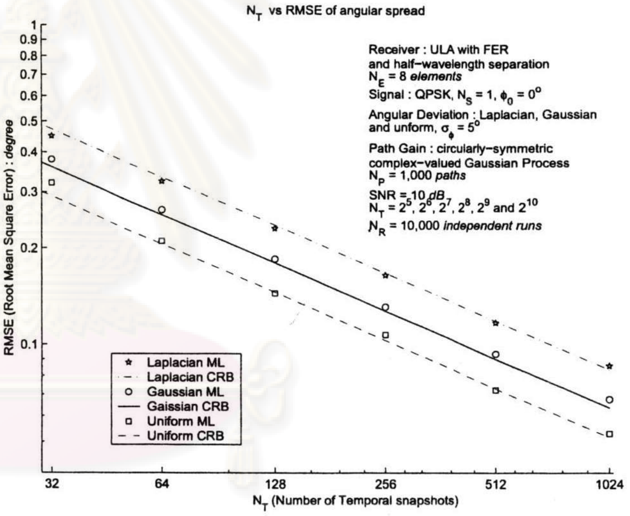


Fig. 2. RMSE of the angular spread σ_ϕ in three individual scenarios with appropriate assumptions.

As seen pictorially in whatever PDF we have correctly assumed, empirical RMSEs of ML estimates in Fig. 1 and 2 agree very well with their theoretical CRB performances attainable from nominal angle and angular spread estimations. The uniform distribution enables CRB to the lowest for all of them in either Fig. 1 or Fig. 2. Intuitively, owing to the fact that both Laplacian and Gaussian angle deviations belong to the same class of symmetrical and exponential distribution, the nominal angle estimation errors in Laplacian deviation is very slightly less than that in Gaussian deviation. It is probably due to the well-known fact that the heavier tail of Laplacian PDF results in higher chance to encounter the large angle deviation *i.e.*, angular outlier. Unlike the nominal angle

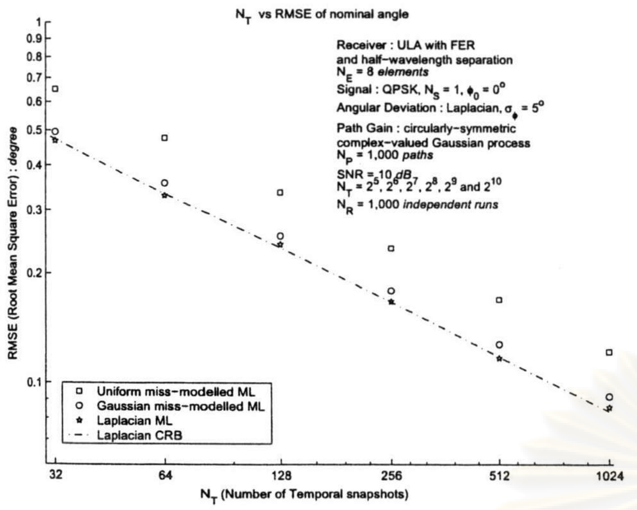


Fig. 3. RMSE of the nominal angle ϕ in Laplacian scenario with correct and incorrect assumptions.

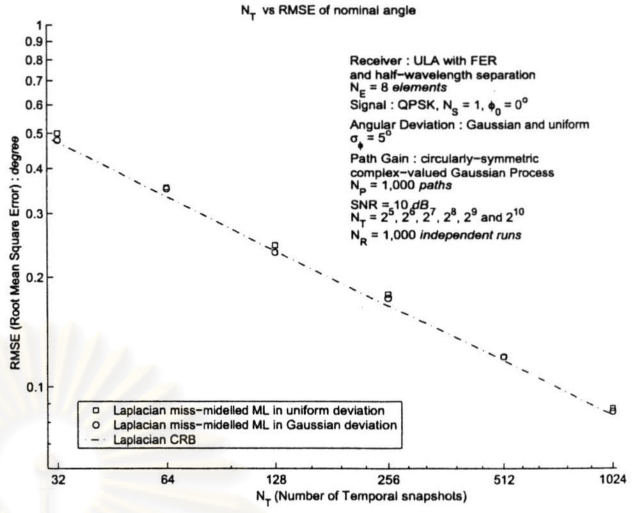


Fig. 5. RMSE of the nominal angle ϕ in two individual scenarios with Laplacian assumption.

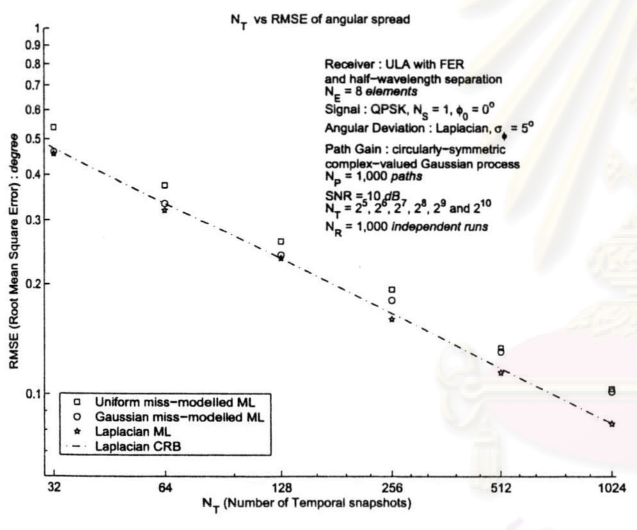


Fig. 4. RMSE of the angular spread σ_ϕ in Laplacian scenario with correct and incorrect assumptions.

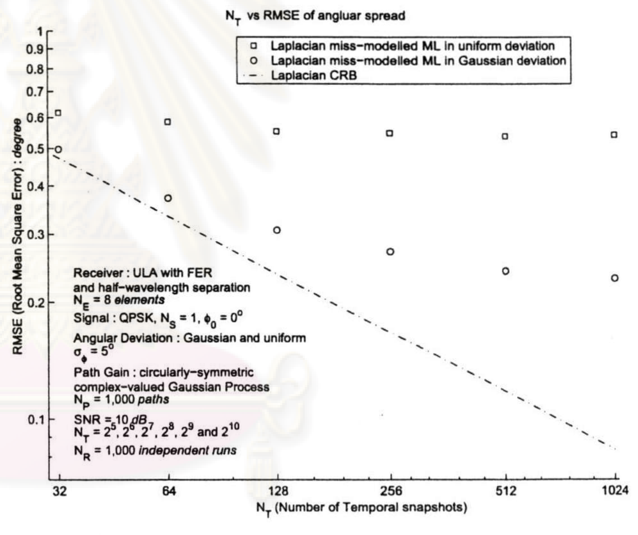


Fig. 6. RMSE of the angular spread σ_ϕ in two individual scenarios with Laplacian assumption.

estimation, angular spread parametrization requires itself more exact assumption relying on $\boldsymbol{B}(\sigma_\phi)$ (see (11) and Laplacian (36)). Therefore, Laplacian RMSE degradation is noticeable in angular spread estimation of Fig. 2.

As stated involving the symmetrically exponential family, miss-modelled estimations with Gaussian assumption in Fig. 3 is very close to that given by the true Laplacian assumption. In addition to the RMSE of miss-modelled estimation with Gaussian assumption, RMSE due to uniform assumption in Fig. 3 more differs from one with exact Laplacian assumption. Under the same reason of angular spread parameterization imposed, the DOA estimate errors of Fig. 4 are gradually considerable in whatever severity of miss-modelling. This is due to more closeness to the true quantity of array covariance matrix in large sample region.

Once modelling two other PDFs as Laplacian in Fig. 5, the underlying ML for estimating the nominal angle is inefficient with respect to the true CRBs (see Fig. 1). Nevertheless, the Laplacian CRB seemed, under the miss-modelled assumption of Fig. 5, attainable in nominal angle estimation. The ML estimator is also inconsistent to find angular spread in Fig. 6.

In addition to the efficiency of ML estimator (profitable in all angle models correctly assumed according to Fig. 1 and 2), it is noteworthy that, for nominal angle estimation, both uniform and Gaussian ML estimators in Laplacian scenario are inefficient with respect to Laplacian CRB (see Fig. 3) whereas the Laplacian CRB under Gaussian and uniform miss-modelled assumptions seemed achievable by the Laplacian ML estimator (see Fig. 5). These two explicit notifications allow us to infer that uniform and Gaussian ML estimators are not endurable to concern the Laplacian angle deviation model while Laplacian ML estimator is beneficially robust to tackle both Gaussian and uniform angle deviation models. Such a deduction is however not true when contrasted with estimating the angular spread in Fig. 4 and 6.

VI. Conclusion

A type of angular deviation in parametric channel model has been proposed and then investigated against two most deviations existed in source localization problem. It is unfortunately conceivable that the Laplacian angle model slightly deteriorates nominal angle estimate error and considerably increases angular spread estimate error since there is more chance to encounter outlier directions compared with Gaussian and uniform PDFs. Our empirical and theoretical simulations of estimating the nominal angle yield ignorable model misleading between Gaussian and Laplacian distributions. The robustness of nominal angle estimation in Laplacian deviation was also illustrated via numerical simulations. In order to coincide with recent field measurements, the multipath angle governed by Laplacian distribution, however, is still expected to be a more appropriate candidate for modelling the realistic scattering channel.

Acknowledgment

This work was supported in part by the Cooperation Project between Department of Electrical Engineering and Private Sector for Research and Development, Chulalongkorn University, Thailand. The authors would like to express an appreciation to Joakim Sorelius for the supplement [9] and also thanks to Mats Bengtsson for clarifying the fading statistic assumption in his local scattering model.

References

[1] O. Besson and P. Stoica, "Decoupled estimation of DOA and angular spread for a spatially distributed source," *IEEE Transactions on Signal Processing*, vol. 48, pp. 1872–1882, Jul. 2000.

[2] K. I. Pedersen, P. E. Mogensen and B. H. Fleury, "A stochastic model of the temporal and azimuthal dispersion seen at the base station in outdoor propagation environments," *IEEE Transactions on Vehicular Technology*, vol. 49, no. 2, pp. 437–447, Mar. 2000.

[3] Q. H. Spencer, B. D. Jeffs, M. A. Jensen and A. L. Swindlehurst, "Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel," *IEEE Journal on Selected Areas in Communications*, vol. 3, pp. 347–360, Mar. 2000.

[4] K. I. Pedersen, P. E. Mogensen and B. H. Fleury, "Power azimuth spectrum in outdoor environments," *IEE Electronics Letters*, vol. 33, pp. 1583–1584, Aug. 1997.

[5] ———, "Spatial channel characterstics in outdoor environments and their impact on BS antenna system performance," *The 48th IEEE Vehicular Technology Conference, 1998*, vol.2, pp. 719–724, May 1998.

[6] J. A. Tsai, R. M. Buehrer and B. D. Woerner, "Spatial fading correlation function of circular antenna arrays with Laplacian energy distribution," *IEEE Communications Letters*, vol. 6, pp. 178–180, May 2002.

[7] L. Schumacher, K. Pedersen and P. Mogensen, "From antenna spacings to theoretical capacities–guidelines for simulating MIMO systems," *The 13th IEEE International symposium on Personal, Indoor and Mobile Radio Communications*, vol. 2, pp. 587–592, Sep. 2002.

[8] H. Boujemaa and S. Marcos, "Joint estimation of direction of arrival and angular spread using the knowledge of the power angular density," *The 13th IEEE International symposium on Personal, Indoor and Mobile Radio Communications*, vol. 4, pp. 1517–1521, Sep. 2002.

[9] J. Sorelius, "Subspace-based parameter estimation problems in signal processing," Ph.D. Dissertation, Department of Information Technology, Uppsala University, Uppsala, Sweden, Feb. 1999.

[10] S. Valaee, "Array processing for detection and localization of narrowband, wideband and distributed sources" Ph. D. Dissertation, Telecommunications and Signal Processing Laboratory, Department of Electrical Engineering, McGill University, Montreal, Canada, May 1994.

[11] T. Trump and B. Ottersten, "Estimation of nominal direction of arrival and angular spread using an array of sensors," Technical Report no. IR-S3-SB-9607, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, Apr. 1996. also appeared in *Signal Processing*, vol. 50, pp. 57–69, Apr. 1996.

[12] P. Zetterberg, "Mobile cellular communications with base station antenna arrays: spectrum efficiency, algorithms and propagation models," Ph.D. Dissertation no. TRITA-S3-SB-9712, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, Jun. 1997.

[13] M. Bengtsson, "Antenna array signal processing for high rank data models," Ph.D. Dissertation no. TRITA-S3-SB-9938, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, Dec. 1999.

[14] B. Ottersten, M. Viberg, P. Stoica and A. Nehorai, "Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing," Technical Report no. TRITA-SB-9302, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, Feb. 1993. also appeared in *Radar Array Signal Processing*, L. Litva, S. Haykin and T. J. Shephred (Eds.), Berlin, Germany: Springer Verlag, ch. 4, pp. 99–151, 1993.

[15] B. Göransson, "On parametric methods for source localization," Ph.D. Dissertation no. TRITA-S3-SB-9727, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, Nov. 1997.

[16] M. Jansson, "On subspace methods in system identification and sensor array signal processing". Ph.D. Dissertation no. TRITA-REG-9701, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, Oct. 1997.

[17] B. Ottersten, P. Stoica and R. Roy, "Covariance matching estimation techniques for array signal processing applications," Technical Report no. IR-S3-SB-9849, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, Jul. 1998. also appeared in *Digital Signal Processing, A Review Journal*, vol. 8, pp. 185–210, Jul. 1998.

[18] P. Stoica, E. G. Larsson and A. B. Gershman, "The stochastic CRB for array processing: a textbook derivation," *IEEE Signal Processing Letters*, vol. 8, pp. 148–150, May 2001.

[19] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[20] N. Johnson, S. Kotz and N. Balakrishnan, *Continuous Univariate Distributions Volume 2*. 2nd edition, NY: John Wiley & Sons, Inc., 1995.

[21] M. Abramowitz and I. A. Stegun (Eds.), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. 9th printing, Dover, NY: 1972.

[22] E. Parzen, *Modern Probability Theory and Its Applications*. NY: John Wiley & Sons, Inc., 1960.

[23] M. Evans, N. Hastings and B. Peacock, *Statistical Distributions*. 2nd edition, NY: John Wiley & Sons, Inc., 1993.

On Toeplitz-Constrained Weights for Spatially Distributed Source Localization

Bamrung Täu Sieskul Kiattisak Maichalernnukul and Somchai Jitapunkul
Digital Signal Processing Research Laboratory, Department of Electrical Engineering, Chulalongkorn University
Bangkok, Thailand, 10330, Tel: +662-0-2218-6900, Fax: +662-0-2218-6912
E-mail: {bamrung.t, kiattisak.m}@student.chula.ac.th, somchai.j@chula.ac.th

***Abstract*— Two contributions are affordable in this paper. Firstly, we provide a relationship between two well-known methods—redudancy averaging (RA) and weighted covariance-matching (WCM)—for estimating any Toeplitz Hermitian covariance matrix. The analysis presented herein enables us to their connection by mean of optimal weight performed. By applying both Toeplitzifications in asymptotic weight determination, another virtue is to propose an improvement on the asymptotic best consistent (ABC) estimator for direction finding in the presence of spatially distributed source. Since the Toeplitz structure of array covariance matrix is held itself in symmetric angle deviation model, it is thus advantageous to incorporate such a priori knowledge of the matrix structure into consistent weight estimation step. Based on weighted least squares (WLS) criteria, our enforcement upon the Toeplitz structure is pictorially illustrated that, in small or even moderate number of temporal snapshots, the estimations of nominal direction and their associating angular spread with the additional constraint can outperform the conventional approach without preprocessing of structured covariance estimation.**

I. Introduction

Sensor array processing plays a prominent role in the phenomenons of the propagation of plane waves transmitted through a media. The problem of finding their directions impinging on array antenna or sensor array, called *direction finding*, is of interest long times ago [1]. This is because it is a useful parameter in several systems such as wireless commination, radar, navigation and *etc*.

In circularly-symmetric complex-valued stochastic process, the multivariate second-order statistic imposes the covariance matrix with Hermitian structure. It is well-known that the sample covariance matrix is an unstructured maximum likelihood estimate of any covariance matrix with Hermitian structure [2]. However, uniform sampling period results in more restriction on the covariance structure. In signal processing application, the use of uniform linear array (ULA) with phase reference at the first element leads to Toeplitz structure which occurs in virous situations [3].

By means of spatially distributed source localization, an objective is to estimate the nominal direction and/or the underlying angular spread whereas the nuisance parameters are such as channel gain, signal power and noise variance. Of particular interest in the spatially distributed source localization problem is the Toeplitz structure in the true quantity of array covariance matrix [4]. This is due to the angle deviation model which is in most drawn from symmetric PDF family [5].

Most works involved the direction finding problem were based on maximum likelihood (ML) estimator due to its producible optimality [1], [6]. In order to arrive at extremal quantity, the optimization search of likelihood function seems inevitable for a complex model. However, the physical model with well-described characterization would, in general, requires large number of model parameters. As a consequence, the larger the number of model parameters, the larger the dimension of optimization search over parameter space. Unfortunately, this might prohibit the ML estimator from being incorporated into real-word applications.

Trying to reduce the optimization tasks, the weighted least squares (WLS) approach is preferred instead of the criterion based on likelihood function [7]. Such an idea stems from the fact that both ML and WLS methods yield the same asymptotic performance when taking into acount any Gauss-Markov model (see *e.g.*, [8, pp. 127–128], [9, pp. 566–567] and [10]). For making the WLS more attractive in computational cost than the ML, there were various efforts to reduce the dimension of optimization during computing the parameter estimate. A reasonable way is to replace the optimal weight with other one which is consistent, or obviously, converges to the optimal weight. The most famous way usually concerns a nonparametric estimation of such value, *e.g.*, the sample estimate. The cause of this is that in a fairly large situation, such a statistic is not only easy to compute but also holds the performance of asymptotic consistency [6].

The ordinary WLS seems, however, inefficient to hold the statistical performance in non-asymptotic region, *i.e.*, a small or even moderate number of samples. Towards this non-asymptotic end, two methods which incorporate the Toeplitz structure have been argued to increase the direction estimate. In the classical (point source) model of sensor array processing, the redundancy averaging (RA) method was investigated to carry out correlated signals [11]. Later, a weight covariance-matching (WCM) approach was proposed to efficiently handle uncorrelated signals [12]. The reason why one would concern these two nonparametric estimates stems from the facts that: i) there exists a closed form for each solution which results in low computational complexity and ii) both of them converge, in probability, to the true covariance.

Here we proceed on the source localization by first investigating the relationship between RA and WCM approaches. The purpose of this is to asses their achievable performances when one desires to make use of both methods. Owing to the fact that the WCM belongs to an asymptotic best consistent (ABC) estimate, the result shown earlier persuades us to replace the ordinary sample weight matrix utilized in [13] with one exploring the Toeplitz structure. This allows us to be higher in performance than that utilized the ordinary sample covariance matrix because the improved weight matrix more approaches the exact and optimal weight matrix even in small number of temporal snapshots. Note that this enhancement does not require any much more computational cost as one expected since the nonparametric weight is conducted only one time before performing the required numerical search.

The arrangement of this paper is as follows. Spatially distributed source localization model is introduced in section II for being demonstrated as an example which is familiar with an incident of Toeplitz structure. Section III reviews the covariance Toeplitzifications through RA and WCM approaches. At the last of the section, it includes the relationship of both methods. We further indicate the application of both approaches to WLS loss function in section IV. With respect to the classical one employing the sample mean covariance, numerical examples are pictorially conducted in section V to validate the improved angle estimates due to both nonparametric weight estimates. In section VI, the results of this paper is finally summarized.

II. Spatially Distributed Source Model

Restrict our attention to a number of source signals, N_S, which are transmitting trough a dispersive channel and then impinging on a sensor array. In the so-called first element reference (FER), the array response vector $\boldsymbol{a}(\phi)$: $[-90°, 90°] \mapsto \mathbb{C}^{N_E \times 1}$ can be, in general, written ideally as

$$\boldsymbol{a}(\phi) \triangleq \begin{bmatrix} 1 & e^{\imath k d_E \sin(\phi)} & \cdots & e^{\imath k d_E (N_E - 1)\sin(\phi)} \end{bmatrix}^{\mathsf{T}} \quad (1)$$

where N_E is the number of sensor elements, d_E signifies the equi-distance between two adjacent elements $k = \frac{2\pi}{\lambda}$ denotes the wave number whose associating wavelength is λ. In flat fading channel, the array output at time instant $n_T \in \{1, 2, \ldots, N_T\}$, or the snapshot vector $\boldsymbol{x}[n_T] \in \mathbb{C}^{N_E \times 1}$, can be represented as [5, p. 25]

$$\boldsymbol{x}[n_T] = \boldsymbol{H}[n_T]\boldsymbol{s}[n_T] + \boldsymbol{n}[n_T] \quad (2)$$

where $\boldsymbol{n}[n_T] \in \mathbb{C}^{N_E \times 1}$ designates the additive noise at sensor array, $\boldsymbol{H}[n_T] \in \mathbb{C}^{N_E \times N_S}$ and $\boldsymbol{s}[n_T] \in \mathbb{C}^{N_S \times 1}$ are respectively constituted as

$$\boldsymbol{H}[n_T] \triangleq \begin{bmatrix} \boldsymbol{h}_1[n_T] & \boldsymbol{h}_2[n_T] & \cdots & \boldsymbol{h}_{N_S}[n_T] \end{bmatrix} \quad (3a)$$

$$\boldsymbol{s}[n_T] \triangleq \begin{bmatrix} s_1[n_T] & s_2[n_T] & \cdots & s_{N_S}[n_T] \end{bmatrix}^{\mathsf{T}} \quad (3b)$$

and N_T denotes the total number of temporal snapshots collected in a data burst $\boldsymbol{X} \in \mathbb{C}^{N_E \times N_T}$ as

$$\boldsymbol{X} \triangleq \begin{bmatrix} \boldsymbol{x}[1] & \boldsymbol{x}[2] & \cdots & \boldsymbol{x}[N_T] \end{bmatrix}. \quad (4)$$

As previously developed, most local scattering models assume that the nominal angle ϕ is deterministic while angular deviation δ_ϕ and associating path gain γ are considered as stochastic quantities. Let N_P be the time-invariant number of scattering paths. For a large number of incoming rays, the channel vector

$$\boldsymbol{h}[n_T] \triangleq \sum_{n_P=1}^{N_P} \gamma_{n_P}[n_T]\boldsymbol{a}(\phi + \delta_{\phi_{n_P}}[n_T]) \quad (5)$$

seemed, under the central limit theorem, plausible to hold a circularly-symmetric complex-valued Gaussian process, *i.e.*, $\boldsymbol{h}[n_T] \sim \mathcal{N}_c(\boldsymbol{0}; \boldsymbol{\Sigma}_{hh}, \boldsymbol{O})$ [14]. This N_E-dimensional variate implicitly provides the statistic $\boldsymbol{\Sigma}_{hh} \triangleq \mathcal{E}\langle \tilde{\boldsymbol{h}}[n_T]\tilde{\boldsymbol{h}}^{\mathsf{H}}[n_T]\rangle \in \mathbb{C}_{\mathbb{H}}^{N_E \times N_E}$, $\tilde{\boldsymbol{h}}[n_T] \triangleq \boldsymbol{h}[n_T] - \mathcal{E}\langle \boldsymbol{h}[n_T]\rangle = \boldsymbol{h}[n_T]$. For taking into account an incoherently distributed channel [15], the second-rder statistic of a certain incoming ray yields [5]

$$\mathcal{E}\langle \gamma_{n_P}[n_T]\gamma^*_{\acute{n}_T}[\acute{n}_T]\rangle = \sigma_\gamma^2 \delta_{n_P,\acute{n}_P}\delta_{n_T,\acute{n}_T} \quad (6)$$

where $\delta_{\bullet,\bullet}$ denotes the Krönecker delta function and σ_γ^2 is the power due to any path. Over the spatial continuum of incoming rays, it can be approximated as

$$\boldsymbol{\Sigma}_{hh}(\rho, \phi, \sigma_\phi) \approx \rho \int f(\delta_\phi|0; \sigma_\phi^2)\boldsymbol{a}(\phi + \delta_\phi)\boldsymbol{a}^{\mathsf{H}}(\phi + \delta_\phi)d\delta_\phi \quad (7)$$

where $\rho \triangleq N_P\sigma_\gamma^2$ signifies the cluster power due to all paths and $f(\delta_\phi|0; \sigma_\phi^2)$ is a conditional PDF accounted for random deviation δ_ϕ given a priori knowledge of the angular spread σ_ϕ. As encountered, a family of symmetric distributions with zero mean and variance σ_ϕ^2 is in most modelled as follows

$$f(\delta_\phi|0; \sigma_\phi^2) = \begin{cases} \frac{1}{2\sqrt{3}\sigma_\phi} \sqcap [-\sqrt{3}\sigma_\phi, \sqrt{3}\sigma_\phi] & ;\text{Uniform} \\ \frac{1}{\sqrt{2\pi}\sigma_\phi} e^{-\frac{1}{2}\frac{\delta_\phi^2}{\sigma_\phi^2}} & ;\text{Gaussian} \\ \frac{1}{\sqrt{2}\sigma_\phi} e^{-\frac{1}{\sigma_\phi}\sqrt{2}|\delta_\phi|} & ;\text{Laplacian.} \end{cases} \quad (8)$$

Rather than such physical angles ϕ and σ_ϕ, the spatial frequency response is preferable due to the better accuracy of approximating the first-order Taylor series around the array broadside [5]. In general, the spatial frequency ω and its associating standard deviation σ_ω are provided by

$$\omega(\phi) = kd_E \sin(\phi) \quad (9a)$$

$$\sigma_\omega(\phi, \sigma_\phi) = kd_E \cos(\phi)\sigma_\phi. \quad (9b)$$

Accounting for small angular spread, the so-called *spatial frequency* approximation results in a separable form as

$$\boldsymbol{\Sigma}_{hh}(\rho, \omega, \sigma_\omega) \simeq \rho \boldsymbol{D}_a(\omega)\boldsymbol{B}(\sigma_\omega)\boldsymbol{D}_a^{\mathsf{H}}(\omega) \quad (10)$$

where $\boldsymbol{D}_a(\omega) : [-kd_E, kd_E] \mapsto \mathbb{C}_{\mathbb{D},\mathbb{U}}^{N_E \times N_E}$ is diagonal and unitary matrix parameterized and $\boldsymbol{B}(\sigma_\omega) : \mathbb{R}_+^{1\times 1} \mapsto \mathbb{R}_{\mathbb{S},\mathbb{T}}^{N_E \times N_E}$ is symmetric Toeplitz matrix. Both of their $(n_E, \acute{n}_E)$-th elements can be represented by [5, p. 22]

$$[\boldsymbol{D}_a(\omega)]_{[n_E,\acute{n}_E]} = e^{\imath(n_E - 1)\omega}\delta_{n_E,\acute{n}_E} \quad (11a)$$

$$[\boldsymbol{B}(\sigma_\omega)]_{[n_E,\acute{n}_E]} = f_{\mathcal{F}}((n_E - \acute{n}_E)\sigma_\omega|0, 1) \quad (11b)$$

whence characteristic function $f_{\mathcal{F}}(t|, 0, 1) \triangleq \mathcal{F}(f(\delta_\omega|, 0, 1))$ is equivalent to the Fourier transform $\mathcal{F}(\cdot)$ of the PDF whose associating random variable holds zero-mean and unit variance. In a certain situation, the $(n_E, \acute{n}_E)$-th element aligned in $B(\sigma_\omega)$ can be expressed as (see *e.g.*, [5] and [16])

$$[B(\sigma_\omega)]_{[n_E,\acute{n}_E]} = \begin{cases} \frac{\sin((n_E-\acute{n}_E)\sqrt{3}\sigma_\omega)}{(n_E-\acute{n}_E)\sqrt{3}\sigma_\omega} & ;\text{uniform} \\ e^{-\frac{1}{2}(n_E-\acute{n}_E)^2\sigma_\omega^2} & ;\text{Gaussian} \\ \frac{1}{1+\frac{1}{2}(n_E-\acute{n}_E)^2\sigma_\omega^2} & ;\text{Laplacian.} \end{cases} \quad (12)$$

If the additive noise assumed is spatially uncorrelated noise and absolutely uncorrelated from channels, it results in

$$\Sigma_{xx}[n_T] = \sum_{n_S=1}^{N_S} p_{n_S}[n_T] D_a(\omega_{n_S}) B(\sigma_{\omega_{n_S}}) D_a^{\mathsf{H}}(\omega_{n_S}) + \sigma_n^2 I \quad (13)$$

where $p_{n_S}[n_T] \triangleq \rho_{n_S} |s_{n_S}[n_T]|^2$ stands for the n_S-th signal power observed at the sensor array. Considering the array covariance matrix in (13), the actual Σ_{xx} is Toeplitz due to $B(\sigma_{\omega_{n_S}})$ being. For more details, we refer the reader to a verification on this quotation in [4].

In what follows, we shall consider only the deterministic signal with constant modulus so that $\Sigma_{xx}[n_T] = \Sigma_{xx}(\theta_\circ)\ ;\forall n_T$, where $\theta_\circ$ is the true value of model parameter. Now assume that, based on the Toeplitz Hermitian covariance matrix $\Sigma_{xx}(\theta_\circ) : \mathbb{R}^{(3N_S+1)\times 1} \mapsto \mathbb{C}_{\mathbf{H},\mathbf{T}}^{N_E\times N_E}$, our problem is to find the nominal direction of arrival, ϕ, given the collected data $x[n_T]; \forall n_T$, where true-valued parameter vector $\theta_\circ \in \mathbb{R}^{(3N_S+1)\times 1}$ defined by

$$\theta_\phi \triangleq \begin{bmatrix} \phi^{\mathsf{T}} & \sigma_\phi^{\mathsf{T}} & p^{\mathsf{T}} & \sigma_n^2 \end{bmatrix}^{\mathsf{T}} \quad (14a)$$

$$\theta_\omega \triangleq \begin{bmatrix} \omega^{\mathsf{T}} & \sigma_\omega^{\mathsf{T}} & p^{\mathsf{T}} & \sigma_n^2 \end{bmatrix}^{\mathsf{T}} \quad (14b)$$

are given respectively for representing the model parameter in term of physical and spatial frequency models.

III. Covariance Toeplitzifications

As already mentioned that $x[n_T] \sim \mathcal{N}_c(0; \Sigma_{xx}, O)$, the sample covariance matrix $\hat{\Sigma}_{xx} \in \mathbb{C}_{\mathbf{H}}^{N_E\times N_E}$ can be given by

$$\hat{\Sigma}_{xx} = \frac{1}{N_T} \sum_{n_T=1}^{N_T} x[n_T] x^{\mathsf{H}}[n_T]. \quad (15)$$

Since $\hat{\Sigma}_{xx} = \Sigma_{xx}(\theta_\circ) + O_p(1/\sqrt{N_T})$, the sample covariance $\hat{\Sigma}_{xx}$ thus converges in large sample to be of Toeplitz Hermitian structure. Notice that the true covariance matrix satisfies $\Sigma_{xx}(\theta_\circ) = \bar{\Sigma}_{xx}(\tau) : \mathbb{R}^{(2N_E-1)\times 1} \mapsto \mathbb{C}_{\mathbf{H},\mathbf{T}}^{N_E\times N_E}$ and then can be expressed in Toeplitz Hermitian form as

$$\bar{\Sigma}_{xx}(\tau) = \begin{bmatrix} \tau_0 & \tau_1^* & \cdots & \tau_{N_E-1}^* \\ \tau_1 & \tau_0 & \ddots & \tau_{N_E-2}^* \\ \vdots & \ddots & \ddots & \vdots \\ \tau_{N_E-1} & \tau_{N_E-2} & \cdots & \tau_0 \end{bmatrix} \quad (16)$$

where the fundamental vector $\tau \in \mathbb{R}^{(2N_E-1)\times 1}$ of all lower triangular entries of the Toeplitz Hermitian matrix Σ_{xx} is

$$\tau \triangleq \begin{bmatrix} \tau_0 & \Re(\tau_1) & \Im(\tau_1) & \cdots & \Re(\tau_{N_E-1}) & \Im(\tau_{N_E-1}) \end{bmatrix}^{\mathsf{T}}. \quad (17)$$

Let $L_{n_L} \in \mathbb{B}^{N_E\times N_E}$ be the block-lower triangular matrix corresponding to the $n_L \in \{0, 1, \ldots, N_E - 1\}$-th lag

$$L_{n_L} \triangleq \begin{bmatrix} O_{(n_L\times(N_E-n_L))} & O_{(n_L\times n_L)} \\ I_{(N_E-n_L)} & O_{((N_E-n_L)\times n_L)} \end{bmatrix}. \quad (18)$$

To review two following methodologies of Toeplitzification, two full-column rank matrices, such as the binary selection matrix $\breve{\Xi} \in \mathbb{B}_{\mathbb{F}}^{N_E^2\times(2N_E-1)}$ and the complex selection matrix $\Upsilon \in \mathbb{C}_{\mathbb{F}}^{(2N_E-1)\times(2N_E-1)}$, are introduced according to

$$\breve{\Xi} \triangleq \begin{bmatrix} v_{\mathsf{c}}(I) & v_{\mathsf{c}}(L_1) & v_{\mathsf{c}}(L_1^{\mathsf{T}}) & \cdots & v_{\mathsf{c}}(L_{N_E-1}) & v_{\mathsf{c}}(L_{N_E-1}^{\mathsf{T}}) \end{bmatrix} \quad (19a)$$

$$\Upsilon \triangleq \begin{bmatrix} 1 & 0^{\mathsf{T}} \\ 0 & I_{(N_E-1)} \otimes \begin{bmatrix} 1 & \imath \\ 1 & -\imath \end{bmatrix} \end{bmatrix}. \quad (19b)$$

where $v_{\mathsf{c}}(\cdot)$ designates the column-stacking vectorization and $\otimes$ signifies the Krönecker product.

A. Redundancy-Averaging Toeplitzification

One of the easiest ways to Toeplitzize the covariance matrix is to observe all entries aligned in a subdiagonal of $\hat{\Sigma}_{xx}$. Averaging all $N_E - n_L$ redundant lags of $[\hat{\Sigma}_{xx}]_{[n_L+n_l, n_l]}$, it attributes the covariance estimate according to *redundancy averaging*. And the averaged result is designated as the n_L-th Toeplitz lag $\hat{\tau}_{n_L} \in \mathbb{C}^{1\times 1}$ computed by [11]

$$\hat{\tau}_{n_L} = \frac{1}{N_E - n_L} \sum_{n_l=1}^{N_E-n_L} [\hat{\Sigma}_{xx}]_{[n_L+n_l, n_l]}. \quad (20)$$

Keeping it into the fundamental vector $\hat{\tau}_{\mathrm{RA}} \in \mathbb{R}^{(2N_E-1)\times 1}$ as

$$\hat{\tau}_{\mathrm{RA}} = \begin{bmatrix} \hat{\tau}_0 & \Re(\hat{\tau}_1) & \Im(\hat{\tau}_1) & \cdots & \Re(\hat{\tau}_{N_E-1}) & \Im(\hat{\tau}_{N_E-1}) \end{bmatrix}^{\mathsf{T}} \quad (21)$$

then the Toeplitz-Hermitian covariance estimate is written as

$$v_{\mathsf{c}}(\hat{\bar{\Sigma}}_{xx}(\hat{\tau}_{\mathrm{RA}})) = \breve{\Xi} \Upsilon \hat{\tau}_{\mathrm{RA}}. \quad (22)$$

B. Weighted Covariance-Matching Toeplitzification

Let $\|\mathbf{A}\|_W^2 \triangleq v_{\mathsf{c}}^{\mathsf{H}}(\mathbf{A}) W^{-1} v_{\mathsf{c}}(\mathbf{A})$ be a weighted version of the Euclidean norm, where $\mathbf{W}$ is a positive-definite Hermitian weight matrix. Based on the extended invariance principle, the WCM can be reformulated as [12]

$$\begin{aligned} \hat{\tau}_{\mathrm{WCM}} &= \arg\min_{\tau} \| \bar{\Sigma}_{xx}(\tau) - \hat{\Sigma}_{xx} \|_W^2 \\ &= \Upsilon^{-1} (\breve{\Xi}^{\mathsf{T}} W^{-1} \breve{\Xi})^{-1} \breve{\Xi}^{\mathsf{T}} W^{-1} \hat{\xi}_x \end{aligned} \quad (23)$$

where $\hat{\xi}_x = v_{\mathsf{c}}(\hat{\Sigma}_{xx}) \in \mathbb{C}^{N_E^2\times 1}$ and $W \in \mathbb{C}_{\mathbb{H}}^{N_E^2\times N_E^2}$ are the covariance vectorization and Hermitian weight, respectively.

To make the residual $\tilde{\boldsymbol{\tau}}_{\text{WCM}} \triangleq \hat{\boldsymbol{\tau}}_{\text{WCM}} - \boldsymbol{\tau}$ minimal, the optimal weight should be satisfied by [7]

$$\boldsymbol{W} = \lim_{N_T \to \infty} N_T \mathcal{E}\langle \tilde{\boldsymbol{\xi}}_x \tilde{\boldsymbol{\xi}}_x^{\mathsf{H}} \rangle = \boldsymbol{\Sigma}_{xx}^{\mathsf{T}} \otimes \boldsymbol{\Sigma}_{xx} \tag{24}$$

where $\tilde{\boldsymbol{\xi}}_x \triangleq \hat{\boldsymbol{\xi}}_x - \boldsymbol{\xi}_x$ is the the sample covariance residual. Since the exact weight depends itself on the model parameter, it is preferable to make use of nonparametric estimate $\hat{\boldsymbol{W}} = \hat{\boldsymbol{\Sigma}}_{xx}^{\mathsf{T}} \otimes \hat{\boldsymbol{\Sigma}}_{xx}$ rather than $\boldsymbol{W}$ without loss of asymptotic performance. Therefore, the WCM covariance estimate $\hat{\bar{\boldsymbol{\Sigma}}}_{xx}(\hat{\boldsymbol{\tau}}_{\text{WCM}})$ can be devectorized from [17]

$$\boldsymbol{v}_{\text{C}}(\hat{\bar{\boldsymbol{\Sigma}}}_{xx}(\hat{\boldsymbol{\tau}}_{\text{WCM}})) \cong \breve{\boldsymbol{\Xi}}(\breve{\boldsymbol{\Xi}}^{\mathsf{T}}\hat{\boldsymbol{W}}^{-1}\breve{\boldsymbol{\Xi}})^{-1}\breve{\boldsymbol{\Xi}}^{\mathsf{T}}\hat{\boldsymbol{W}}^{-1}\hat{\boldsymbol{\xi}}_x. \tag{25}$$

It is however fruitful to see the relation between the WCM and RA estimate. Furthermore, if it needs to deploy one of both estimates, *i.e.*, the WCM and RA fundamental vector estimates, their statistical properties should be concerned.

Lemma 1 (reformable weight of WCM to be RA): Let the optimal weight $\boldsymbol{W}$ *in (23) be the identity matrix. Then, a consistent fundamental vector estimate* $\hat{\boldsymbol{\tau}}_{\text{CM}} \in \mathbb{R}^{(2N_E-1)\times 1}$ *which is given by*

$$\begin{aligned}\hat{\boldsymbol{\tau}}_{\text{CM}} &\triangleq \arg\min_{\boldsymbol{\tau}} \|\bar{\boldsymbol{\Sigma}}_{xx}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Sigma}}_{xx}\|_{\boldsymbol{I}}^2 \\ &= \boldsymbol{\Upsilon}^{-1}(\breve{\boldsymbol{\Xi}}^{\mathsf{T}}\breve{\boldsymbol{\Xi}})^{-1}\breve{\boldsymbol{\Xi}}^{\mathsf{T}}\hat{\boldsymbol{\xi}}_x\end{aligned} \tag{26}$$

is equivalent to that available from the RA method i.e.,

$$\hat{\boldsymbol{\tau}}_{\text{CM}} = \hat{\boldsymbol{\tau}}_{\text{RA}}. \tag{27}$$

Proof: See appendix I. ■

Notice that (22) might be rewritten as

$$\boldsymbol{v}_{\text{C}}(\hat{\bar{\boldsymbol{\Sigma}}}_{xx}(\hat{\boldsymbol{\tau}}_{\text{RA}})) = \boldsymbol{\Pi}_{\mathcal{S}_{\mathcal{R}}(\breve{\boldsymbol{\Xi}})}\hat{\boldsymbol{\tau}}_{\text{RA}} \tag{28}$$

where $\boldsymbol{\Pi}_{\mathcal{S}_{\mathcal{R}}(\mathbf{A})} = \mathbf{A}(\mathbf{A}^{\mathsf{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{H}}$ signifies the orthogonal projection onto the range space of any full rank matrix $\mathbf{A}$. Such a lemma is intended to illustrate the similarity between unweighed WCM approach and RA method. Based on best linear unbiased estimator (BLUE) viewpoint [10], the performance of WCM must be, in principle, superior than that available from the RA method because the first one invoked the optimal weight to reduce the residual.

IV. Weighted Least Squares Estimation

The WLS estimation is to find a parametric argument which provides the smallest residual in matching criteria. In spatially distributed source localization, most work is relied upon the second-order statistics. Therefore, the array covariance matrix is one to be matched between theoretical and receivable quantities [13]. Being concentrated on the source signal power and noise variance, the WLS loss function is arrived at

$$\hat{\boldsymbol{\vartheta}}_{\text{WLS}} = \arg\min_{\boldsymbol{\vartheta}_\omega} \hat{\boldsymbol{\xi}}_x^{\mathsf{H}} \boldsymbol{W}^{-\frac{1}{2}} \boldsymbol{\Pi}^{\perp}_{\mathcal{S}_{\mathcal{R}}(\boldsymbol{W}^{-\frac{1}{2}}\boldsymbol{\Omega}(\boldsymbol{\vartheta}_\omega))} \boldsymbol{W}^{-\frac{1}{2}} \hat{\boldsymbol{\xi}}_x \tag{29}$$

where $\boldsymbol{\vartheta}_\omega \triangleq [\boldsymbol{\omega}^{\mathsf{T}} \ \ \boldsymbol{\sigma}_\omega^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{R}^{(2N_S)\times 1}$ results in $\boldsymbol{\Omega}(\boldsymbol{\vartheta}_\omega)$ expressed by

$$\begin{aligned}\boldsymbol{\Omega}(\boldsymbol{\omega}, \boldsymbol{\sigma}_\omega) \triangleq \Big[\, &\boldsymbol{v}_{\text{C}}(\boldsymbol{D}_a(\omega_1)\boldsymbol{B}(\sigma_{\omega_1})\boldsymbol{D}_a^{\mathsf{H}}(\omega_1)) \quad \cdots \\ &\boldsymbol{v}_{\text{C}}(\boldsymbol{D}_a(\omega_{N_S})\boldsymbol{B}(\sigma_{\omega_{N_S}})\boldsymbol{D}_a^{\mathsf{H}}(\omega_{N_S})) \quad \boldsymbol{v}_{\text{C}}(\boldsymbol{I}) \,\Big].\end{aligned} \tag{30}$$

To deal with single signal scenario, the parameter $\boldsymbol{\vartheta}$ will be reduced to ω and its corresponding $\boldsymbol{\Omega}(\omega)$ is given from [4]

$$\boldsymbol{\Omega}(\omega) \triangleq \left(\boldsymbol{D}_a^{\mathsf{H}}(\omega) \otimes \boldsymbol{D}_a(\omega)\right) \boldsymbol{\Xi} \tag{31}$$

where the binary selection matrix $\boldsymbol{\Xi} \in \mathbb{B}_{\mathbb{F}}^{N_E^2 \times N_E}$ is

$$\boldsymbol{\Xi} \triangleq \left[\boldsymbol{v}_{\text{C}}(\boldsymbol{I}) \quad \boldsymbol{v}_{\text{C}}(\boldsymbol{L}_1 + \boldsymbol{L}_1^{\mathsf{T}}) \quad \cdots \quad \boldsymbol{v}_{\text{C}}(\boldsymbol{L}_{N_E-1} + \boldsymbol{L}_{N_E-1}^{\mathsf{T}})\right]. \tag{32}$$

For terminology concise, one may recall $\hat{\boldsymbol{\vartheta}}_{\text{WLS}}$ as *ordinary WLS* estimate when replacing $\boldsymbol{W}$ in (29) with $\hat{\boldsymbol{W}}$ in (24) and (15).

Proposition 1: By substituting $\boldsymbol{W}$ in (29) with $\hat{\bar{\boldsymbol{W}}}(\hat{\boldsymbol{\tau}}_{\text{RA}}) \triangleq \hat{\bar{\boldsymbol{\Sigma}}}_{xx}^{\mathsf{T}}(\hat{\boldsymbol{\tau}}_{\text{RA}}) \otimes \hat{\bar{\boldsymbol{\Sigma}}}_{xx}(\hat{\boldsymbol{\tau}}_{\text{RA}})$ and $\hat{\bar{\boldsymbol{W}}}(\hat{\boldsymbol{\tau}}_{\text{WCM}}) \triangleq \hat{\bar{\boldsymbol{\Sigma}}}_{xx}^{\mathsf{T}}(\hat{\boldsymbol{\tau}}_{\text{WCM}}) \otimes \hat{\bar{\boldsymbol{\Sigma}}}_{xx}(\hat{\boldsymbol{\tau}}_{\text{WCM}})$, we shall designate the improved solution as *RA-WLS* and *WCM-WLS* estimates, respectively.

V. Numerical Examples

In addition to verification of the lemma 1, the following simulations also demonstrate the impact of two proposed improvements. In all experiments setting up, we commonly employ the ULA with half-wavelength separation to receive a QPSK (quaternary phase shift keying) signals whose strength are controllable with respect to noise variance by SNR $\triangleq 10\log\left(\sigma_s^2 / \sigma_n^2\right)$. All significant parameters are set up, unless otherwise a variation on the parameter of interest will be specified individually in each figure, as the following table:

ϕ_o	σ_{ϕ_o}	σ_γ^2	SNR	N_P	N_E
$-10°, 0°, 10°$	5°	0.01	10	100	8

Practically, the pseudo random number satisfied the Laplacian PDF $f_{\text{L}}(\delta_\phi|0,1)$ can be modified from [16]

$$\delta_{\phi_{\text{L}}} = \frac{1}{\sqrt{2}} \ln\left(\frac{\delta_{\phi_{\text{U}}}}{\acute{\delta}_{\phi_{\text{U}}}}\right) \tag{33}$$

with any two independent uniform distributions $\delta_{\phi_{\text{U}}} \sim \mathcal{U}[0,1]$ and $\acute{\delta}_{\phi_{\text{U}}} \sim \mathcal{U}[0,1]$. Our empirical standard deviation (RMSE) is calculated from a large number of independent runs (N_R).

By inspecting the single source case in Fig. 1, 2 and 3, all WLS-based estimators asymptotically achieve the corresponding CRBs, *i.e.*, uniform, Gaussian and Laplacian CRBs, respectively. It is noteworthy that the WLS estimators employing the Toeplitz-constrained weights, such as RA and WCM covariance estimates, outperform the ordinary WLS, particularly in small and moderate numbers of temporal snapshots. Furthermore, both converge to the CRB at hand more rapid than the ordinary WLS. This is due to the fact that, in each improvement, the consistent weight matrix has been forced beforehand to be of Toeplitz structure accounting for a property of optimal weight. For more insight into the improved weight reflections, the WCM-WLS is better than the RA-WLS, especially in small number of temporal snapshots. This superiority agrees well with the lemma 1 which can be infered that the WCM weight estimate will provide error less than that given from RA weight estimation.

In multiple source scenario, the uniform and Gaussian angle deviation model is of interest in Fig. 4. Their angular

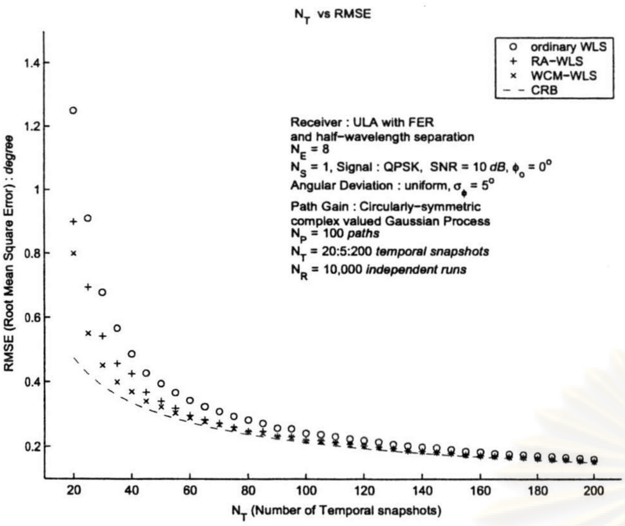


Fig. 1. Uniform angle deviation: empirical and theoretical standard deviations of the errors due to estimating the nominal angle ϕ as a function of the number of temporal snapshots N_T.

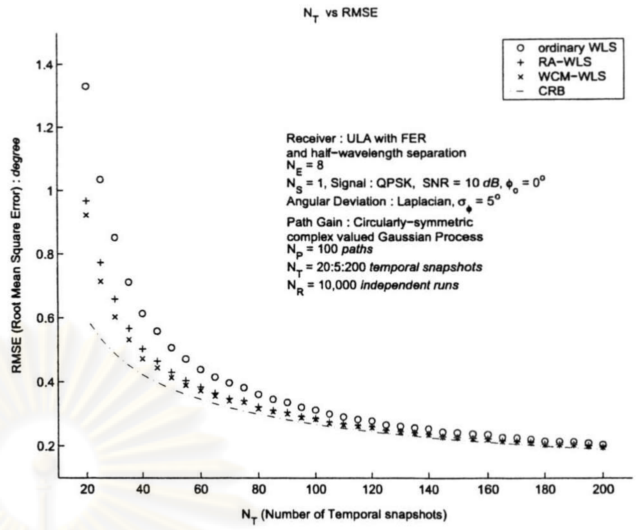


Fig. 3. Laplacian angle deviation: empirical and theoretical standard deviations of the errors due to estimating the nominal angle ϕ as a function of the number of temporal snapshots N_T.

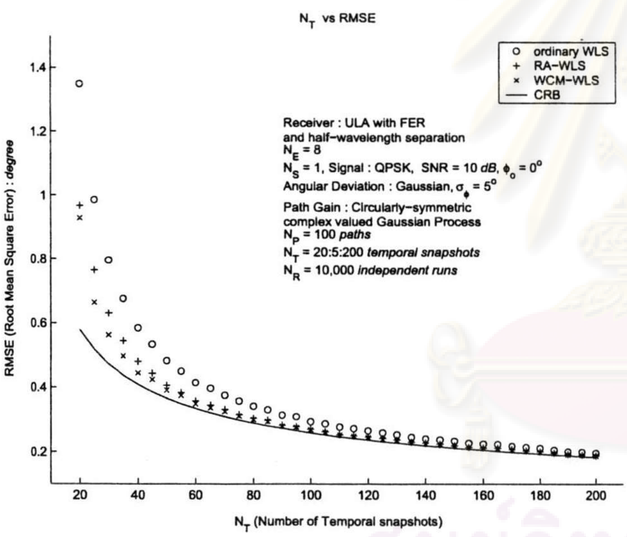


Fig. 2. Gaussian angle deviation: empirical and theoretical standard deviations of the errors due to estimating the nominal angle ϕ as a function of the number of temporal snapshots N_T.

separation and angular spreads are fixed as 20 and 5 degrees, respectively. For the uniform angle deviation cluster, maximum angle according to (8) becomes $-10 + \sqrt{3} \times 5 \approx -1.3397$ degree with respect to array broadside. In the Gaussian angle source, this leads to small spatial ambiguity because an arrival direction in the cluster yields rather low probability to align in another cluster. Although the Gaussian CDF giving $\int_{-\infty}^{-1.3397} f(\delta_\phi|10;5^2)d\delta_\phi \approx 0.0117$ is of low probability, such a correlation is evidently observable when investigating the unattainable gaps to uniform deviation CRBs in both nominal direction and angular spread estimations at small number of temporal snapshots. In Gaussian case, there exist violent degradations of unweighed criteria at small number of temporal snapshot, where the RA-WLS might be more worse than the ordinary WLS. Indeed, this region demonstrates a fluctuation suppression due to employing the ABC weight $\hat{\boldsymbol{W}}$ in (25) for preparing the WCM-WLS weight $\hat{\hat{\boldsymbol{W}}}(\hat{\boldsymbol{\tau}}_{\text{WCM}})$.

VI. Conclusion

A variational expression between the redundancy averaging [11] and weighted covariance-matching [12] is first explored. It is indicated that the RA methodology is equivalent to the unweighed version of the WCM loss function. According to the principle of best linear unbiased estimator, the WCM estimate with optimal weight outperforms that given from the RA method in any Gauss-Makov model. Such a statement is then verifiable when incorporating both Toeplitzifications into the spatially distributed source localization. Replacing the consistent weight matrix due to sample covariance estimate with one due to RA or WCM covariance estimate, numerical examples also illustrate that the RMSEs of estimated directions can be improved over the ordinary weighted least squares criteria in both single and multiple source cases [4] and [13]. As seen pictorially, the RMSEs of WLS invoked RA and WCM agree well with the relationship indicated herein.

Appendix I
Proof of the Lemma 1

The solution (26) is easily derived by replacing $\boldsymbol{W}$ in (23) with $\boldsymbol{I}$. To verify (27), we reformulate the residual norm as

$$\begin{aligned}\hat{\boldsymbol{\tau}}_{\text{CM}} &= \arg\min_{\boldsymbol{\tau}} \|\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Sigma}}_{xx}\|_{\text{F}}^2 \\ &= \arg\min_{\boldsymbol{\tau}} \lceil (\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Sigma}}_{xx})^2 \rfloor.\end{aligned} \tag{34}$$

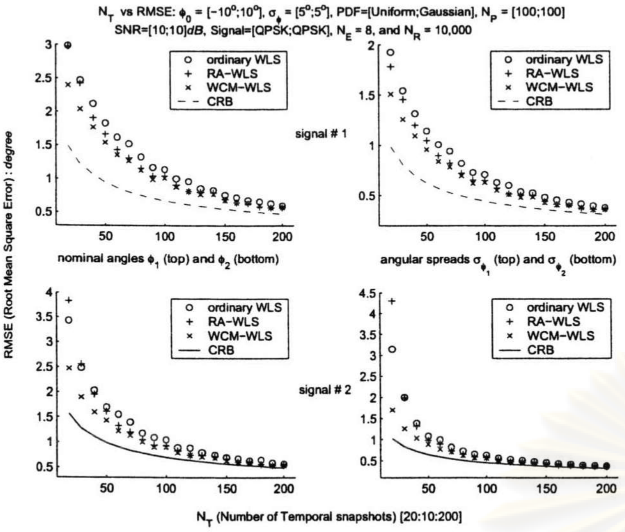


Fig. 4. Uniform and Gaussian angle deviations: empirical and theoretical standard deviations of the estimated direction errors as a function of the number of temporal snapshots N_T.

where $\|\mathbf{A}\|_{\mathrm{F}}^2 \triangleq \lceil \mathbf{A}^{\mathrm{H}}\mathbf{A} \rfloor$ and $\lceil \mathbf{A} \rfloor$ denote the Frobenius norm and matrix trace respectively. Since the analytic function of CM needs to be satisfied by $\frac{\partial}{\partial \tau_{n_L}^*}\lceil(\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Sigma}}_{xx})^2\rfloor = 0\,;\forall n_L$ [18, p. 891], its complex-valued derivative with respect to the n_L-th Toeplitz lag results in

$$\begin{aligned}&\frac{\partial}{\partial \tau_{n_L}^*}\lceil(\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Sigma}}_{xx})^2\rfloor \\ &= -2\lceil \hat{\boldsymbol{\Sigma}}_{xx}\dot{\boldsymbol{\Sigma}}_{xx}(\tau_{n_L}^*)\rfloor + 2\lceil \boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau})\dot{\boldsymbol{\Sigma}}_{xx}(\tau_{n_L}^*)\rfloor\end{aligned} \tag{35}$$

where $\dot{\mathbf{A}}(\mathrm{x}) \triangleq \frac{\partial}{\partial \mathrm{x}}\mathbf{A}(\mathrm{x})$ is the derivative with respect to x. Let us represent $\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau})$ in a linear structure according to

$$\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) = \tau_0 \boldsymbol{I} + \sum_{n_L=1}^{N_E-1} \tau_{n_L}\boldsymbol{L}_{n_L} + \tau_{n_L}^*\boldsymbol{L}_{n_L}^{\mathsf{T}}. \tag{36}$$

Then, it follows that

$$\dot{\boldsymbol{\Sigma}}_{xx}(\tau_{n_L}^*) = \boldsymbol{L}_{n_L}^{\mathsf{T}}. \tag{37}$$

Forcing $\frac{\partial}{\partial \tau_{n_L}^*}\lceil(\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Sigma}}_{xx})^2\rfloor \trianglerighteq 0$, we obtain the critical condition of the n_L-th Toeplitz lag as

$$\lceil \boldsymbol{L}_{n_L}\hat{\boldsymbol{\Sigma}}_{xx}\rfloor = \lceil \boldsymbol{L}_{n_L}\boldsymbol{\Sigma}_{xx}(\boldsymbol{\tau})\rfloor. \tag{38}$$

Proceeding on the n_L-th subdiagonal, it yields

$$\hat{\tau}_{n_L} = \frac{1}{N_E - n_L}\sum_{n_l=1}^{N_E-n_L}[\hat{\boldsymbol{\Sigma}}_{xx}]_{[n_L+n_l,n_l]} \tag{39}$$

which coincides with one available from the RA lag in (20).

ACKNOWLEDGMENT

This work was supported in part by the Cooperation Project Between Department of Electrical Engineering and Private Sector for Research and Development, Chulalongkorn University, Thailand. The authors would like to express an appreciation to all collaborators' hospitalities during the numerical simulations.

REFERENCES

[1] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, pp. 97–94, Jul. 1996.

[2] T. W. Anderson, *The Statistical Analysis of Time Series*. NY: John Wiley & Sons, Inc., 1971.

[3] J. P. Burg, D. G. Luenberger, and D. L. Wenger, "Estimation of structured covariance matrices," *Proceedings of the IEEE*, vol. 70, pp. 963–974, Sept. 1982.

[4] O. Besson and P. Stoica, "Decoupled estimation of DOA and angular spread for a spatially distributed source," *IEEE Transactions on Signal Processing*, vol. 48, pp. 1872–1882, Jul. 2000.

[5] M. Bengtsson, "Antenna array signal processing for high rank data models," Ph.D. Dissertation no. TRITA-S3-SB-9938, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, Dec. 1999.

[6] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing," Technical Report no. TRITA-SB-9302, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, Feb. 1993. also appeared in *Radar Array Signal Processing,* L. Litva, S. Haykin and T. J. Shephred (Eds.), Berlin, Germany: Springer Verlag, ch. 4, pp. 99–151, 1993.

[7] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," Technical Report no. IR-S3-SB-9849, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, Jul. 1998. also appeared in *Digital Signal Processing, A Review Journal,* vol. 8, pp. 185–210, Jul. 1998.

[8] R. J. Muirhead, *Aspect of Multivariate Statistical Theory.* NY: John Wiley & Sons, 1982.

[9] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. 2nd edition, NY: John Wiley & Sons, 1984.

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

[11] D. A. Linebarger and D. H. Johnson, "The effect of spatial averaging on spatial correlation matrices in the presence of coherent signals," *IEEE Transactions on Signal Processing*, vol. 38, pp. 880–884, May 1990.

[12] H. Li, P. Stoica, and J. Li, "Computationally efficient maximum likelihood estimation of structured covariance matrices," *IEEE Transactions on Signal Processing*, vol. 47, pp. 1314–1323, May 1999.

[13] T. Trump and B. Ottersten, "Estimation of nominal direction of arrival and angular spread using an array of sensors," Technical Report no. IR-S3-SB-9607, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, Apr. 1996. also appeared in *Signal Processing*, vol. 50, pp. 57–69, Apr. 1996.

[14] P. Zetterberg, "Mobile cellular communications with base station antenna arrays: spectrum efficiency, algorithms and propagation models," Ph.D. Dissertation no. TRITA-S3-SB-9712, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, Jun. 1997.

[15] S. Valaee, "Array processing for detection and localization of narrowband, wideband and distributed sources" Ph. D. Dissertation, Telecommunications and Signal Processing Laboratory, Department of Electrical Engineering, McGill University, Montreal, Canada, May 1994.

[16] B. Täu Sieskul and S. Jitapunkul, "Towards Laplacian angle deviation model for spatially distributed source localization," to be published in *International Symposium on Communications and Information Technologies 2004 (ISCIT 2004)*, Aug. 2004.

[17] D. A. Turkington, *Matrix Calculus and Zero-One Matrices: Statistical and Econometric Applications.* Cambridge University Press, 2002.

[18] S. Haykin, *Adaptive Filter Theory.* 3rd edition, Englewood Cliffs, NJ: Prentice-Hall, 1996.

Towards Incorporation of Toeplitz Constraint into Asymptotic Maximum Likelihood for Estimating Nominal Direction of Spatially Distributed Source*

B. Täu Sieskul S. Jitapunkul
Center of Excellence in Telecommunication Technology, Department of Electrical Engineering
Faculty of Engineering, Chulalongkorn University, Bangkok, 10200, Thailand
Phone +662-0-2218-6915, Fax +662-0-2218-6912, E-mail bts@chula.com, somchai.j@chula.ac.th

Abstract

This paper is intended to address the covariance matrix estimation and to improve the estimation of nominal direction in Rayleigh channel owing to spatially distributed source by employing a sensor array receiver. Three contributions are affordable herein. Firstly, we propose a large-sample approximation of maximum likelihood criterion for estimating a covariance matrix with linearly affine structure. Later, a connection of the proposed estimator to an existed criterion is secondly provided. It is conceivable that the presented criterion yields the same solution as derived from the weighted covariance-matching method. The key idea behind application area stems from noticing that array covariance matrix is not only Hermitian but also Toeplitz. Based on such priori knowledge, it is advantageous to make use of structured covariance matrix in parameter estimation. To decrease the directional estimate error, the last contribution is to incorporate the imposed Toeplitz-Hermitian matrix into another large-sample approximated maximum likelihood estimator. Numerical simulations are conducted to verify and indicate that, particularly in non-asymptotic region, the asymptotically efficient estimator with Toeplitz constraint always outperforms one loosing the additional information.

Keywords: Toeplitz-Hermitan structure, sensor array processing, parameter estimation

1 Introduction

Sensor array processing deals, in general, with inferring the parameter of source signal based on the use of uniform linear array (ULA) [1]. With phase reference at the first element, this enables array covariance matrix to Toeplitz structure. More precisely, the signal and the additive noise must also be stochastically uncorrelated in the classical (point source) model. As taken into account the multipath directions, the classical model is further argued to be unrealistic. However, the array covariance matrix still be possible to hold Toeplitz structure. The reason why Toeplitz structure is held as well belongs to; firstly, sensor elements are aligned as uniformly linear and phase-referenced at the first element, another one is due to the symmetry of random deviation around nominal direction.

Towards this end, all elements in array covariance should be forced to make the estimated covariance matrix Hermitian and Toeplitz. Due to approaching the exact value more rapid than the ordinary sample covariance, the estimated covariance matrix with Toeplitz constraint seems advantageous when having to replace the parametric array covariance matrix with a non-parametric estimate.

Here our purpose is to proceed on two problems, such as, estimation of matrix with linearly affine structure which can be seen by a special case as Toeplitz structure, and spatially distributed source localization. A large-sample approximation of maximum likelihood for estimating the structured matrix is proposed herein. We then provide a connection to other approach. Systematically, a lemma is set up in order to verify that the proposed estimator accounting for structured matrix yields the same solution being available from the weighted covariance-matching criterion [2]. During the non-parametric array covariance estimation, we provide another contribution to improve the asymptotic maximum likelihood (AML) estimator proposed in [3] by invoking the Toeplitz restriction. This modification does not need any much more additional calculation because the Toeplitz-constrained covariance estimate belongs to non-parametric type which can be computed only one time before the optimization search.

2 Spatially Distributed Source Model

Restrict our attention to the source signal transmitting through dispersive channel and then impinging on the sensor array. With first element reference (FER), the array response vector $\boldsymbol{a}(\phi) : [-\frac{\pi}{2}, \frac{\pi}{2}] \mapsto \mathbb{C}^{N_E \times 1}$ is written ideally as

$$\boldsymbol{a}(\phi) \triangleq \begin{bmatrix} 1 & e^{\imath k d_E \sin(\phi)} & \dots & e^{\imath k d_E (N_E - 1)\sin(\phi)} \end{bmatrix}^{\mathsf{T}} \quad (1)$$

*This work was supported in part by the Cooperation Project between Department of Electrical Engineering and Private Sector for Research and Development, Chulalongkorn University, Thailand.

where $k = \frac{2\pi}{\lambda}$ designates the wave number with associating wavelength λ, d_E signifies the equi-distance between two adjacent elements, and N_E is the number of sensor elements. In flat-fading channel, the array output at $n_T \in \{1, 2, \ldots, N_T\}$—the snapshot vector $\boldsymbol{x}[n_T] \in \mathbb{C}^{N_E \times 1}$— can be represented as [4, p. 25]

$$\boldsymbol{x}[n_T] = s[n_T] \sum_{n_P=1}^{N_P} \gamma_{n_P}[n_T]\boldsymbol{a}(\phi + \delta_{\phi,n_P}[n_T]) + \boldsymbol{n}[n_T] \tag{2}$$

where $s[n_T] \in \mathbb{C}^{1\times 1}$, $\gamma_{n_P}[n_T] \in \mathbb{C}^{1\times 1}$, $\phi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, δ_{ϕ,n_P}, $\boldsymbol{n}[n_T] \in \mathbb{C}^{N_E \times 1}$ and N_P signify complex baseband signal, path gain, nominal direction, angle deviation, additive noise at sensor array and number of multipath directions, consecutively[1]. In what follows, our interest belongs to constant modulus signal, *i.e.*, $|s[n_T]|^2$ is time-invariant. Let all stochastic quantities are all uncorrelated with each other. Based on second-order statistic, the path gain and the noise vector are characterized such that $\gamma_{n_P}[n_T] \sim \mathcal{N}_c(0; \sigma_\gamma^2, 0)$ [5] and $\boldsymbol{n}[n_T] \sim \mathcal{N}_c(\boldsymbol{0}; \sigma_n^2 \boldsymbol{I}, \boldsymbol{O})$, where the first and last zeros are due to circularly-symmetric complex-valued stochastic processes with zero mean. Under the central limit theorem, the snapshot vector enables us to $\boldsymbol{\Sigma}_{xx} \triangleq \mathcal{E}\langle \tilde{\boldsymbol{x}}[n_T]\tilde{\boldsymbol{x}}^{\mathsf{H}}[n_T]\rangle \in \mathbb{C}_{\mathbb{H}}^{N_E \times N_E}$ which is derived from $\tilde{\boldsymbol{x}}[n_T] = \boldsymbol{x}[n_T] - \mathcal{E}\langle \boldsymbol{x}[n_T]\rangle$. In spatial frequency model, two directions are parameterized as

$$\omega(\phi) = kd_E \sin(\phi) \tag{3a}$$

$$\sigma_\omega(\phi, \sigma_\phi) = kd_E \cos(\phi)\sigma_\phi. \tag{3b}$$

In term of such a reparameterization, the array covariance matrix might be rewritten as [4]

$$\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta}) = p\boldsymbol{D}_a(\omega)\boldsymbol{B}(\sigma_\omega)\boldsymbol{D}_a^{\mathsf{H}}(\omega) + \sigma_n^2 \boldsymbol{I} \tag{4}$$

where $p = N_P \sigma_\gamma^2 |s[n_T]|^2$ is the observable power and the model parameter $\boldsymbol{\theta} \in \mathbb{R}^{4\times 1}$ in spatial frequency representation, the diagonal unitary matrix $\boldsymbol{D}_a(\omega) : [-kd_E, kd_E] \mapsto \mathbb{C}_{\mathbb{D},\mathbb{U}}^{N_E \times N_E}$ parameterized by nominal angle, and the symmetric Toeplitz matrix $\boldsymbol{B}(\sigma_\omega) : \mathbb{R}_+^{1\times 1} \mapsto \mathbb{R}_{\mathbb{S},\mathbb{T}}^{N_E \times N_E}$ parameterized by angular spread matrix, can be expressed as

$$\boldsymbol{\theta}_\omega = \begin{bmatrix} \omega & \sigma_\omega & p & \sigma_n^2 \end{bmatrix}^{\mathsf{T}} \tag{5a}$$

$$[\boldsymbol{D}_a(\omega)]_{[n_E, \acute{n}_E]} = e^{\imath(n_E - 1)\omega}\delta_{n_E, \acute{n}_E} \tag{5b}$$

$$[\boldsymbol{B}(\sigma_\omega)]_{[n_E, \acute{n}_E]} = f_{\mathcal{F}}((n_E - \acute{n}_E)\sigma_\omega | 0, 1). \tag{5c}$$

And the characteristic function $f_{\mathcal{F}}(t|, 0, 1) \triangleq \mathcal{F}(f(\delta_\omega|, 0, 1))$ is equivalent to the Fourier transform $\mathcal{F}(\bullet)$ of the associating PDF whose random variable holds zero-mean and unit variance. In a certain situation, the $(n_E, \acute{n}_E)$-th element in $\boldsymbol{B}(\sigma_\omega)$ can be expressed as (see *e.g.*, [4] and [6])

$$[\boldsymbol{B}(\sigma_\omega)]_{[n_E, \acute{n}_E]} = \begin{cases} \frac{\sin((n_E - \acute{n}_E)\sqrt{3}\sigma_\omega)}{(n_E - \acute{n}_E)\sqrt{3}\sigma_\omega} & \text{; uniform} \\ e^{-\frac{1}{2}(n_E - \acute{n}_E)^2\sigma_\omega^2} & \text{; Gaussian} \\ \frac{1}{1 + \frac{1}{2}(n_E - \acute{n}_E)^2\sigma_\omega^2} & \text{; Laplacian.} \end{cases} \tag{6}$$

As arrived at $\boldsymbol{x}[n_T] \sim \mathcal{N}_c(\boldsymbol{0}; \boldsymbol{\Sigma}_{xx}, \boldsymbol{O})$, the sample mean covariance $\hat{\boldsymbol{\Sigma}}_{xx} \in \mathbb{C}_{\mathbb{H}}^{N_E \times N_E}$ is computed from

$$\hat{\boldsymbol{\Sigma}}_{xx} = \frac{1}{N_T} \sum_{n_T=1}^{N_T} \boldsymbol{x}[n_T]\boldsymbol{x}^{\mathsf{H}}[n_T]. \tag{7}$$

3 Toeplitz Covariance Estimation

When there is no a priori knowledge of Toeplitz structure, it is conceivable that the sample covariance $\hat{\boldsymbol{\Sigma}}_{xx}$ in (7) is the unstructured maximum likelihood estimate of $\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta})$ [7]. Since the array covariance matrix is itself Toeplitz, one should be aware of such a matrix structure during a parameter estimation.

In this section, we derive an array covariance Toeplitizification and then consider its relationship to other loss function. The way to explore the Toeplitz structure in $\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta})$ can be given by rewriting as $\bar{\boldsymbol{\Sigma}}_{xx}(\boldsymbol{\tau}) : \mathbb{R}^{(2N_E-1)\times 1} \mapsto \mathbb{C}_{\mathbb{T},\mathbb{H}}^{N_E \times N_E}$, where $\boldsymbol{\tau} \in \mathbb{R}^{(2N_E-1)\times 1}$ denotes the fundamental vector of Toeplitz-Hermitian matrix $\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta})$ (see *e.g.*, [3] for more details). In the form of linearly affine structure, we can write[2]

$$\bar{\boldsymbol{\xi}}_x(\boldsymbol{\tau}) = \breve{\boldsymbol{\Xi}}\boldsymbol{\Upsilon}\boldsymbol{\tau} \tag{11}$$

where $\bar{\boldsymbol{\xi}}_x(\boldsymbol{\tau}) \triangleq \boldsymbol{v}_{\mathsf{c}}(\bar{\boldsymbol{\Sigma}}_{xx}(\boldsymbol{\tau})) : \mathbb{R}^{(2N_E-1)\times 1} \mapsto \mathbb{C}^{N_E^2 \times 1}$ designates the array covariance vectorization with $\boldsymbol{v}_{\mathsf{c}}(\cdot)$ designating the column-stacking vectorization operator. Note that $\frac{\partial}{\partial \boldsymbol{\tau}^{\mathsf{T}}}\bar{\boldsymbol{\xi}}_x(\boldsymbol{\tau}) = \breve{\boldsymbol{\Xi}}\boldsymbol{\Upsilon}$. Let $\lceil \cdot \rfloor$ and $|\cdot|$ be the matrix trace and determinant, respectively.

[1] All of these quantities are, in most, classified into two groups, such as deterministic and stochastic processes. For the first kind, the nominal direction ϕ and the complex base-band signal $s[n_T]$ are regarded herein to be deterministic. But another one is to assume that all of the rest are stochastic.

[2] Let us introduce the binary selection of Toeplitz-Hermitian structure as the full column-rank matrix $\breve{\boldsymbol{\Xi}} \in \mathbb{B}_{\mathbb{F}}^{N_E^2 \times (2N_E-1)}$:

$$\breve{\boldsymbol{\Xi}} \triangleq \begin{bmatrix} \boldsymbol{v}_{\mathsf{c}}(\boldsymbol{I}) & \boldsymbol{v}_{\mathsf{c}}(\boldsymbol{L}_1) & \boldsymbol{v}_{\mathsf{c}}^{\mathsf{T}}(\boldsymbol{L}_1) & \cdots & \boldsymbol{v}_{\mathsf{c}}(\boldsymbol{L}_{N_E-1}) & \boldsymbol{v}_{\mathsf{c}}^{\mathsf{T}}(\boldsymbol{L}_{N_E-1}) \end{bmatrix} \tag{8}$$

where $\boldsymbol{L}_{n_L} \in \mathbb{B}^{N_E \times N_E}$ is a block-lower triangular matrix according to

$$\boldsymbol{L}_{n_L} \triangleq \begin{bmatrix} \boldsymbol{O}_{(n_L \times (N_E - n_L))} & \boldsymbol{O}_{(n_L)} \\ \boldsymbol{I}_{(N_E - n_L)} & \boldsymbol{O}_{((N_E - n_L) \times n_L)} \end{bmatrix}. \tag{9}$$

And the full-rank matrix $\boldsymbol{\Upsilon} \in \mathbb{C}_{\mathbb{F}}^{(2N_E-1)\times(2N_E-1)}$ is a complex-valued selection matrix, defined as

$$\boldsymbol{\Upsilon} \triangleq \begin{bmatrix} 1 & \boldsymbol{0}^{\mathsf{T}} \\ \boldsymbol{0} & \boldsymbol{I}_{(N_E-1)} \otimes \begin{bmatrix} 1 & \imath \\ 1 & -\imath \end{bmatrix} \end{bmatrix} \tag{10}$$

where $\otimes$ signifies the Krönecker product operator.

In an unstructured parameterization, the ML estimate and negative normalized likelihood are represented by

$$\hat{\tau}_{\mathrm{ML}} = \arg\min_{\tau} \ell_{\mathrm{ML}}^{[N_T]}(\tau) \tag{12a}$$

$$\ell_{\mathrm{ML}}^{[N_T]}(\tau) \triangleq \lceil \bar{\Sigma}_{xx}^{-1}(\tau)\hat{\Sigma}_{xx} \rfloor + \ln|\bar{\Sigma}_{xx}(\tau)|. \tag{12b}$$

Differentiating (12b) with respect to τ, it yields

$$\frac{\partial}{\partial \tau^{\mathsf{T}}} \ell_{\mathrm{ML}}^{[N_T]}(\tau) = (\breve{\Xi}\Upsilon\tau - \hat{\xi}_x)^{\mathsf{H}} \bar{\Psi}_{xx}^{-1}(\tau)\breve{\Xi}\Upsilon \tag{13}$$

where $\bar{\Psi}_{xx} \triangleq \bar{\Sigma}_{xx}^{\mathsf{T}} \otimes \bar{\Sigma}_{xx} \in \mathbb{C}^{N_E^2 \times N_E^2}$, and $\hat{\xi}_x = v_c(\hat{\Sigma}_{xx})$. Forcing $\frac{\partial}{\partial \tau}\ell_{\mathrm{ML}}^{[N_T]}(\tau) \trianglerighteq \boldsymbol{0}$, we obtain

$$\tau = \left(\Upsilon^{\mathsf{H}}\breve{\Xi}^{\mathsf{T}}\bar{\Psi}_{xx}^{-1}(\tau)\breve{\Xi}\Upsilon\right)^{-1}\Upsilon^{\mathsf{H}}\breve{\Xi}^{\mathsf{T}}\bar{\Psi}_{xx}^{-1}(\tau)\hat{\xi}_x. \tag{14}$$

In what follows, we shall designate the Toeplitz-constrained covariance estimate as $\hat{\hat{\Sigma}}_{xx}(\hat{\tau}_{\mathrm{ML}})$.

Proposition 1. *Since $\hat{\Psi}_{xx} \triangleq \hat{\Sigma}_{xx}^{\mathsf{T}} \otimes \hat{\Sigma}_{xx} \in \mathbb{C}^{N_E^2 \times N_E^2}$ converges in probability to $\bar{\Psi}_{xx}(\tau)$, we can replace the parametric $\bar{\Psi}_{xx}(\tau)$ with the non-parametric $\hat{\Psi}_{xx}$, without loss of asymptotic performance. Relying on* (14), *the covariance estimate is approximated as*

$$v_c(\hat{\hat{\Sigma}}_{xx}(\hat{\tau}_{AML})) = \breve{\Xi}\Upsilon\left(\Upsilon^{\mathsf{H}}\breve{\Xi}^{\mathsf{T}}\hat{\Psi}_{xx}^{-1}\breve{\Xi}\Upsilon\right)^{-1}\Upsilon^{\mathsf{H}}\breve{\Xi}^{\mathsf{T}}\hat{\Psi}_{xx}^{-1}\hat{\xi}_x \tag{15}$$

where AML stands for the attribution of asymptotic maximum likelihood criterion.

We then recall the weighted covariance-matching (WCM) criterion [8]. Let a weighted least squares (WLS) criterion be [2]

$$f_{\mathrm{WLS}}(\tau|W) \triangleq \|\bar{\Sigma}_{xx}(\tau) - \hat{\Sigma}_{xx}\|_W^2 \tag{16}$$

where $\|\mathbf{A}\|_W^2 \triangleq v_c^{\mathsf{H}}(\mathbf{A})W^{-1}v_c(\mathbf{A})$ designates a weighted version of the Euclidean norm. By replacing the consistent weight matrix $\hat{W}$ as $\hat{\Psi}_{xx}$, the WCM estimate of Toepltiz lag is given from

$$\hat{\tau}_{\mathrm{WCM}} = \arg\min_{\tau} f_{\mathrm{WLS}}(\tau|\hat{W}). \tag{17}$$

Lemma 1. *Based on the weighted covariance-matching criterion* (16), *the WCM estimate yields the same solution as provided by the AML estimate, i.e.,*

$$\hat{\tau}_{WCM} = \hat{\tau}_{AML}. \tag{18}$$

Proof. See appendix. □

4 A Large-Sample Approximation of Maximum Likelihood Estimator

To estimate nominal direction of spatially distributed source, in [3] a comparative study of asymptotically efficient estimators is shown that the large-sample approximated maximum likelihood estimator outperforms the WLS-based estimator [8] in non-asymptotic region. Here we try to incorporate the Toeplitz-constrained covariance estimated from (15) into the asymptotic maximum likelihood estimator [3].

Let $\check{\Sigma}_{xx}(\omega) : [-kd_E, kd_E] \mapsto \mathbb{C}_{\mathbb{H}}^{N_E \times N_E}$ be a concentrated covariance matrix which is calculated from

$$v_c(\check{\Sigma}_{xx}(\omega)) = \Omega(\omega)\left(\Omega^{\mathsf{H}}(\omega)\widehat{\Psi}_{xx}^{-1}\Omega(\omega)\right)^{-1}\Omega^{\mathsf{H}}(\omega)\widehat{\Psi}_{xx}^{-1}\hat{\xi}_x \tag{19}$$

where $\widehat{\Psi}_{xx} \in \mathbb{C}_{\mathbb{H}}^{N_E^2 \times N_E^2}$ is arbitrary non-parametric estimate of $\Psi_{xx}(\theta) \triangleq \Sigma_{xx}^{\mathsf{T}}(\theta) \otimes \Sigma_{xx}(\theta)$. Then, the nominal direction estimate in this way is given by

$$\hat{\omega}_{\mathrm{AML}} = \arg\min_{\omega} \ell_{\mathrm{AML}}^{[N_T]}(\omega) \tag{20a}$$

$$\ell_{\mathrm{AML}}^{[N_T]}(\omega) = \lceil \check{\Sigma}_{xx}^{-1}(\omega)\hat{\Sigma}_{xx} \rfloor + \ln|\check{\Sigma}_{xx}(\omega)|. \tag{20b}$$

Letting $\widehat{\Sigma}_{xx}$ in (19) be $\hat{\Sigma}_{xx}$ derived from (7), the above $\hat{\omega}_{\mathrm{AML}}$ owing to (20) will be called the "*ordinary AML*" estimate [3].

Proposition 2. *By invoking $\widehat{\Sigma}_{xx} = \hat{\hat{\Sigma}}_{xx}(\hat{\tau}_{AML})$ according to* (15) *and* (19), *we shall designate the solution $\hat{\omega}_{AML}$ in* (20) *as the estimate of "AML with Toeplitz constraint".*

5 Numerical Examples

For all situations, we assume that the path power in each cluster is normalized as $\rho \triangleq N_P\sigma_\gamma^2 = 1$. Pseudo random number satisfied by the standard Laplacian PDF can be modified from [6].

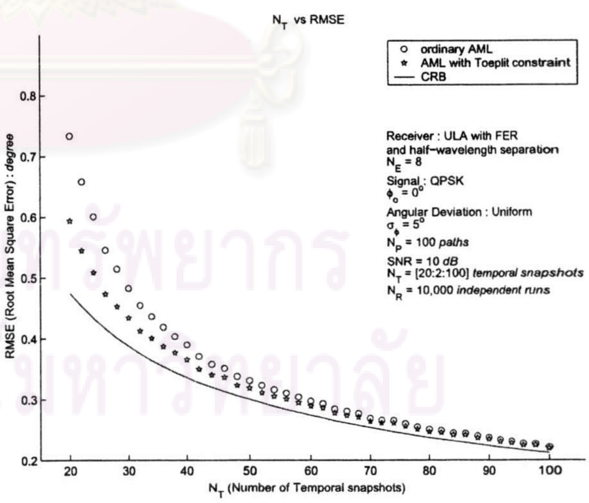


Figure 1: Uniform angle deviation.

In Fig. 1, 2 and 3, we plot empirical and theoretical standard deviations of the errors due to estimating the nominal angle ϕ as a function of number of snapshots N_T. Both AML-based estimates attain each Cramér-Rao bound (CRB) at hand asymptotically, *i.e.*, as the number of temporal snapshot tends to be infinity. Since the AML with Toeplitz constraint

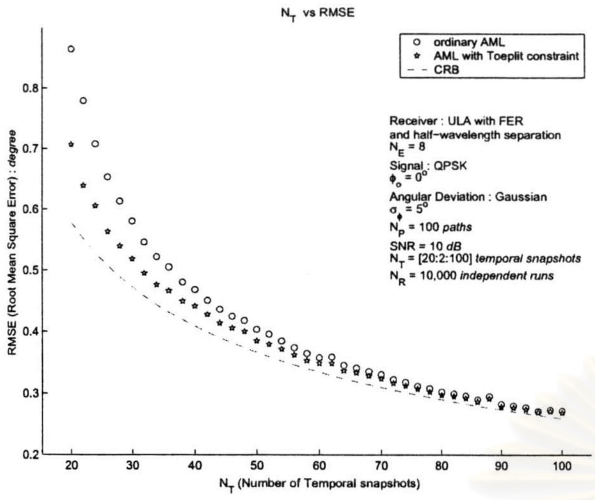


Figure 2: Gaussian angle deviation.

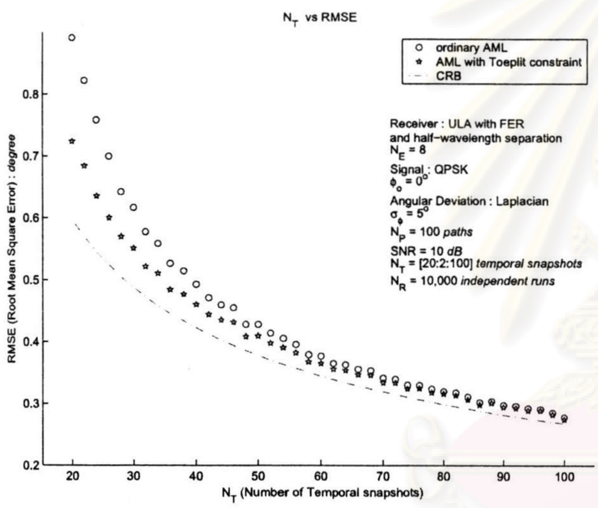


Figure 3: Laplacian angle deviation.

is more accurate than one without this knowledge, the RMSE in this way is thus lower than such ordinary AML. This can be observed in all angular PDFs.

6 Conclusions

An estimation of the matrix with linearly affine structure is proposed and then verified that it yields the same solution of the previous WLS approach. This enables us to a deep insight into the relationship between AML and WLS. For estimating the nominal direction, an improvement of spatially distributed source has been proposed. This contribution stems from the Toeplitz structure imposed in array covariance matrix. Replacing the ordinary sample covariance matrix with the Toeplitz-structured covariance estimate in the large-sample approximated maximum likelihood estimator [3], numerical examples emphasize that it yields better in non-asymptotic performance than one losing this additional knowledge.

Appendix (Proof of Lemma 1)

Inserting (11) into (16), the WLS function is

$$f_{\mathtt{WLS}}(\tau|\hat{\boldsymbol{W}}) = \tilde{\boldsymbol{\xi}}_x^{\mathsf{H}} \hat{\boldsymbol{\Psi}}_{xx}^{-1} \tilde{\boldsymbol{\xi}}_x \qquad (21)$$

where $\tilde{\boldsymbol{\xi}}_x \triangleq \hat{\boldsymbol{\xi}}_x - \breve{\boldsymbol{\Xi}}\boldsymbol{\Upsilon}\boldsymbol{\tau}$. Invoking the chain rule of $\frac{\partial}{\partial \tau^{\mathsf{T}}} f_{\mathtt{WLS}}(\tau|\hat{\boldsymbol{W}}) = \frac{\partial}{\partial \tilde{\boldsymbol{\xi}}_x^{\mathsf{H}}} f_{\mathtt{WLS}}(\tau|\hat{\boldsymbol{W}}) \frac{\partial}{\partial \tau^{\mathsf{T}}} \tilde{\boldsymbol{\xi}}_x$, it results in

$$\frac{\partial}{\partial \tau^{\mathsf{T}}} f_{\mathtt{WLS}}(\tau|\hat{\boldsymbol{W}}) = (2\tilde{\boldsymbol{\xi}}_x^{\mathsf{H}} \hat{\boldsymbol{\Psi}}_{xx}^{-1})(-\breve{\boldsymbol{\Xi}}\boldsymbol{\Upsilon}). \qquad (22)$$

Forcing $\frac{\partial}{\partial \tau} f_{\mathtt{WLS}}(\tau|\hat{\boldsymbol{W}}) \trianglerighteq \boldsymbol{0}$, we obtain

$$\hat{\tau}_{\mathtt{WCM}} = (\boldsymbol{\Upsilon}^{\mathsf{H}} \breve{\boldsymbol{\Xi}}^{\mathsf{T}} \hat{\boldsymbol{\Psi}}_{xx}^{-1} \breve{\boldsymbol{\Xi}} \boldsymbol{\Upsilon})^{-1} \boldsymbol{\Upsilon}^{\mathsf{H}} \breve{\boldsymbol{\Xi}}^{\mathsf{T}} \hat{\boldsymbol{\Psi}}_{xx}^{-1} \hat{\boldsymbol{\xi}}_x \qquad (23)$$

which coincides, through (11), with (15)

References

[1] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Processing Magazine,* vol. 13, pp. 67–94, Jul. 1996.

[2] H. Li, P. Stoica, and J. Li, "Computationally efficient maximum likelihood estimation of structured covariance matrices," *IEEE Transactions on Signal Processing*, vol. 47, pp. 1314–1323, May 1999.

[3] B. Täu Sieskul, "Source direction estimation based on separable parameterization," Master Degree Thesis (in preparation), Chulalongkorn University, Bangkok, Thailand, Aug. 2004.

[4] M. Bengtsson, "Antenna array signal processing for high rank data models," Ph.D. Thesis no. TRITA-S3-SB-9938, Royal Institute of Technology, Stockholm, Sweden, Dec. 1999.

[5] P. Zetterberg, "Mobile cellular communications with base station antenna arrays: spectrum efficiency, algorithms and propagation models," Ph.D. Thesis no. TRITA-S3-SB-9712, Royal Institute of Technology, Stockholm, Sweden, Jun. 1997.

[6] B. Täu Sieskul and S. Jitapunkul, "Towards Laplacian angle deviation model for spatially distributed source localization," accepted to participate in *International Symposium on Communications and Information Technologies 2004 (ISCIT 2004).*

[7] T. W. Anderson, *The Statistical Analysis of Time Series.* NY: John Wiley & Sons, Inc., 1971.

[8] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digital Signal Processing, A Review Journal,* vol. 8, pp. 185–210, Jul. 1998.

Biography

Biography

Bamrung Täu Sieskul was born in Lopburi, Thailand 1979. He received the PMT (Pre Mechanical Technician) certificate from King Mongkut's Institute of Technology North Bangkok, Bangkok, Thailand in 1998, and the B.Eng. degree in Electrical Engineering from Chulalongkorn University, Bangkok, Thailand in 2002. At the moment, he is completing the M.Eng. degree in Electrical Engineering at the same university.

Contributions

1. **B. Täu Sieskul**, J. Pattanavichate and S. Jitapunkul, "Coherent source localization via a spatial smoothing with temporal correlation," *Proceedings of The 1-st Electrical Engineering/Electronics, Computer, Telecommunications, and Information Technology Annual Conference (ECTI-CON 2004)*, pp. 173−176, May 2004.

2. **B. Täu Sieskul**, and S. Jitapunkul, "Decoupled estimation of nominal direction and angular spread based on asymptotic maximum likelihood approach," accepted to participate in *Proceedings of The 4-th WSEAS International Conference on SIGNAL, SPEECH and IMAGE PROCESSING (WSEAS ICOSSIP 2003)*, Sep. 2004.

3. **B. Täu Sieskul** and S. Jitapunkul, "Towards Laplacian angle deviation model for spatially distributed source localization," accepted to participate in *Proceedings of The International Symposium on Communications and Information Technologies 2004 (ISCIT 2004)*, Oct. 2004.

4. **B. Täu Sieskul**, K. Maichalernnukul, and S. Jitapunkul, "On Toeplitz-constrained weights for spatially distributed source localization," accepted to participate in *Proceedings of The International Symposium on Communications and Information Technologies 2004 (ISCIT 2004)*, Oct. 2004.

5. **B. Täu Sieskul** and S. Jitapunkul, "Towards incorporation of Toeplitz Constraint into asymptotic maximum likelihood for estimating nominal direction of spatially distributed source," accepted to participate in *Proceedings of The 27-th Electrical Engineering Conference (EECON 27)*, Nov. 2004.